Supplemental Material: Continuous symmetries and approximate quantum error correction

Philippe Faist,[1,*] Sepehr Nezami,[2,*] Victor V. Albert,[1,3] Grant Salton,[1,2] Fernando Pastawski,[4] Patrick Hayden,[2] and John Preskill[1,3]

[1] *Institute for Quantum Information and Matter, Caltech, Pasadena, CA, USA*
[2] *Stanford Institute for Theoretical Physics, Stanford University, Stanford, CA, USA*
[3] *Walter Burke Institute for Theoretical Physics, Caltech, Pasadena, CA, USA*
[4] *Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Berlin, Germany*

(Dated: July 27, 2020)

Contents:

SM. A: Proof of our bounds for a $U(1)$-covariant code

This section is devoted to the proof of Theorem II and Corollary III. We introduce some auxiliary definitions and recall the corresponding assumptions and statements here for convenience.

Throughout this section we use the following notation. Let L be the logical system of finite dimension d_L, and let $A = A_1 \otimes \cdots \otimes A_n$ be the physical system of finite dimension. Let $\mathcal{E}_{L\to A}$ be any completely positive, trace-preserving map. Consider logical and physical observables T_L and T_A. Let K be a set of subsets of $\{1,\ldots,n\}$. Let $\{q_\alpha\}_{\alpha\in K}$ be a probability distribution and let the noise channel $\mathcal{N}_{A\to AC}$ be given by (11), an expression which we recall here:

$$\mathcal{N}_{A\to AC}(\cdot) = \sum_{\alpha\in K} q_\alpha |\alpha\rangle\langle\alpha|_C \otimes \mathcal{N}^\alpha_{A\to A}(\cdot) \;; \qquad \mathcal{N}^\alpha_{A\to A}(\cdot) = |\phi_\alpha\rangle\langle\phi_\alpha|_{A_\alpha} \otimes \mathrm{tr}_{A_\alpha}(\cdot) \;, \qquad \text{(A.1)}$$

---

[*] These authors contributed equally to this work.

where $A_\alpha = \bigotimes_{i\in\alpha} A_i$ and where $\{|\phi_\alpha\rangle\}_{\alpha\in K}$ are some fixed states.

Furthermore we make use of the following quantities.

Definition 1. Let A be a Hermitian operator on a Hilbert space of dimension d.

(i) The *spectral range of A*, which we denote by ΔA, is the difference between the maximum and the minimum eigenvalue of a Hermitian operator A.

(ii) A *median eigenvalue of A* is a number $\mu \in \mathbb{R}$ such that the length-d vector of eigenvalues of A counted with multiplicity has at least $\lceil d/2\rceil$ components that are less than or equal to μ, and at least $\lceil d/2\rceil$ components that are greater than or equal to μ.

(iii) The *spread of eigenvalues of A around the median*, denoted by $s(A)$, is defined as $s(A) = d^{-1}\|A - \mu\mathbb{1}\|_1$, where μ is a median eigenvalue of A.

(iv) The *spread of eigenvalues of A around the mean*, denoted by $s'(A)$, is defined as $s'(A) = d^{-1}\|A - \mathrm{tr}(A)\mathbb{1}/d\|_1$.

We refer to Propositions 25 and 27 below for an alternative characterization of these quantities, and for a proof that $s(A)$ does not depend on the choice of μ. The median eigenvalue has the same interpretation as the median in statistics: If d is odd, then a median eigenvalue μ is "the middle eigenvalue" of A. If d is even, then μ may be any number in the closed interval separating the two "middle eigenvalues."

Definition 2. We further make use of the following definitions:

(a) The mapping $\mathcal{E}$ is *charge-conserving with respect to T_L and T_A* if there exists $\nu\in\mathbb{R}$ such that $\mathcal{E}^\dagger(T_A) = T_L - \nu\mathbb{1}_L$. Similarly, for $\delta \geqslant 0$ the mapping $\mathcal{E}$ is *δ-charge-conserving with respect to T_L and T_A* if there exists a $\nu\in\mathbb{R}$ such that $\|(T_L - \nu\mathbb{1}_L) - \mathcal{E}^\dagger(T_A)\|_\infty \leqslant \delta$;

(b) Let K' be a set of subsets of $\{1,\ldots n\}$. The charge T_A is *K'-local* if it is of the form $T_A = \sum_{\alpha\in K'} T_\alpha$, where T_α is supported on A_α;

(c) Let $\{t_\alpha^\pm\}_{\alpha\in K'}$, $t_\alpha^\pm \in \mathbb{R}$, $t_\alpha^- \leqslant t_\alpha^+$, and let $\eta \geqslant 0$. Then a K'-local charge T_A is *$(\{t_\alpha^\pm\}_{\alpha\in K'}, \eta)$-bounded for $\mathcal{E}$* if for any logical state σ_L we have

$$\left|\mathrm{tr}\Big(\sum (T_\alpha - t_\alpha)\Pi_\alpha^\perp\, \mathcal{E}(\sigma_L)\Big)\right| \leqslant \eta\ , \tag{A.2}$$

where $\Pi_\alpha^\perp$ projects onto the eigenspaces of T_α whose eigenvalues do not lie in the interval $[t_\alpha^-, t_\alpha^+]$, and where $t_\alpha = (t_\alpha^- + t_\alpha^+)/2$;

Theorem II. *Suppose that $\mathcal{E}$ is charge-conserving with respect to T_L and T_A. Let $K' \subset K$ and assume that T_A is K'-local. Suppose that $q_\alpha > 0$ for all $\alpha\in K'$. Then:*

$$\left.\begin{matrix}\epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E}) \\ \big\langle \epsilon_{\mathrm{e}}(\mathcal{N}^\alpha\circ\mathcal{E})\big\rangle_\alpha\end{matrix}\right\} \geqslant \frac{\max\big\{s(T_L), s'(T_L)/2\big\}}{\max_{\alpha\in K'}(\Delta T_\alpha/q_\alpha)}\ ; \tag{A.3a}$$

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \frac{\Delta T_L/2}{\max_{\alpha\in K'}(\Delta T_\alpha/q_\alpha)}\ , \tag{A.3b}$$

where $\langle\cdot\rangle_\alpha = \sum_{\alpha\in K} q_\alpha(\cdot)$. *(Proof on page 8.)*

Corollary III. *Let $\delta \geqslant 0$ such that the map $\mathcal{E}$ is δ-charge-conserving with respect to T_L and T_A. Let $K' \subset K$ and assume that T_A is K'-local. Suppose that $q_\alpha > 0$ for all $\alpha \in K'$. Let $\{t_\alpha^\pm\}_{\alpha\in K'}$, $t_\alpha^\pm \in \mathbb{R}$, $t_\alpha^- \leqslant t_\alpha^+$, and let $\eta \geqslant 0$ be such that T_A is $(\{t_\alpha^\pm\}_{\alpha\in K'}, \eta)$-bounded for $\mathcal{E}$. Then:*

$$\left.\begin{array}{r}\epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E}) \\ \big\langle \epsilon_{\mathrm{e}}(\mathcal{N}^\alpha\circ\mathcal{E})\big\rangle_\alpha\end{array}\right\} \geqslant \frac{\max\big\{s(T_L), s'(T_L)/2\big\} - \delta - \eta}{\max_{\alpha\in K'}(\Delta t_\alpha/q_\alpha)} \ ; \tag{A.4a}$$

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \frac{\Delta T_L/2 - \delta - \eta}{\max_{\alpha\in K'}(\Delta t_\alpha/q_\alpha)} \ , \tag{A.4b}$$

where $\Delta t_\alpha = t_\alpha^+ - t_\alpha^-$ and where $\langle\cdot\rangle_\alpha = \sum_{\alpha\in K} q_\alpha(\cdot)$. *(Proof on page 10.)*

The proof of Theorem II is split into three lemmas. A first lemma deduces that the environment has access to the logical charge, to a good approximation. The second lemma uses this fact to derive the bounds stated in Theorem II in a more general setting and for any input state. The third lemma specializes these bounds to the setting of the theorem. We then provide a proof of Corollary III with the help of yet another lemma.

Lemma 3. *Suppose that $\mathcal{E}$ is charge-conserving with respect to T_L and T_A and assume that T_A is K'-local for $K' \subset K$. Assume that $q_\alpha > 0$ for all $\alpha \in K'$. Then there exists an observable $Z_{C'E}$ such that*

$$\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(\mathbb{1}_F \otimes Z_{C'E}) = T_L \ ; \qquad\qquad \Delta Z_{C'E} \leqslant \max_\alpha \frac{\Delta T_\alpha}{q_\alpha} \ , \tag{A.5}$$

where the complementary channel $\widehat{\mathcal{N}\circ\mathcal{E}}_{L\to C'EF}$ to the combined encoding and noise is given by (15).

Proof of Lemma 3. First, we may assume without loss of generality that $\|T_\alpha\|_\infty = \Delta T_\alpha/2$, thanks to because we may freely shift $T_\alpha \to T_\alpha + c\mathbb{1}$ without neither impacting the assumptions nor the conclusions of the lemma, and we may therefore arrange for the minimum eigenvalue and the maximum eigenvalue of T_α to differ only by a sign (Proposition 25). Since $\mathcal{E}$ is charge-conserving, let $\nu \in \mathbb{R}$ such that $\mathcal{E}^\dagger(T_A) = T_L - \nu\mathbb{1}$. Again, we may assume without loss of generality that $\nu = 0$ because we may shift $T_L \to T_L + \nu\mathbb{1}$ without changes to the assumptions or the conclusions of the lemma: Any shift in T_L by a multiple of the identity can be reflected by a corresponding shift of $Z_{C'E}$ by the same multiple of the identity, given that $\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger$ is unital. Define the Hermitian operator

$$Z_{C'E} = \sum_{\alpha\in K'} |\alpha\rangle\langle\alpha|_{C'} \otimes (q_\alpha^{-1}\, T_\alpha) \ . \tag{A.6}$$

For any logical state σ_L, and writing $\rho_A = \mathcal{E}(\sigma_L)$,

$$\mathrm{tr}\big(Z_{C'E}\,\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_L)\big) = \sum \mathrm{tr}\big(T_\alpha\ \mathrm{tr}_{A\setminus A_\alpha}(\rho_A)\big) = \mathrm{tr}(T_A\rho_A) = \mathrm{tr}\big(T_L\sigma_L\big) \ . \tag{A.7}$$

Because this relation holds for all states σ_L on L, whose linear span covers all Hermitian operators on L, we have

$$\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(\mathbb{1}_F \otimes Z_{C'E}) = T_L \ . \tag{A.8}$$

Then $\Delta Z_{C'E} \leqslant 2\|Z_{C'E}\|_\infty \leqslant 2\max_{\alpha\in K'}\|T_\alpha\|_\infty = \max_{\alpha\in K'}(q_\alpha^{-1}\Delta T_\alpha)$, since the infinity norm picks out the largest eigenvalue in absolute value. ∎

The second part of the proof of Theorem II is to deduce from the environment's access to the global charge that the code performs poorly with respect to the various entanglement fidelity measures. We phrase this statement as a more general lemma that applies in fact to any noise model, and can be used to bound the fixed-input entanglement fidelity for any given fixed input state $|\phi\rangle_{LR}$, as long as the environment has access to an observable which yields some information about the logical state. In analogy with ϵ_{e} and $\epsilon_{\mathrm{worst}}$, we define for any $|\phi\rangle_{LR}$ and for any channel $\mathcal{N}'$,

$$\epsilon_{|\phi\rangle}(\mathcal{N}') = \sqrt{1 - F^2_{|\phi\rangle}(\mathcal{N}', \mathrm{id})}\ . \tag{A.9}$$

This lemma can be seen as a refinement of Bény's characterization of approximate error correction using operator algebras [103]. To formulate the lemma, we define two auxiliary quantities that depend on a state σ and an observable T:

$$C_{\sigma,T} = \min_{\mu\in\mathbb{R}} \left\| \sigma^{1/2} \left(T - \mu \mathbb{1} \right) \sigma^{1/2} \right\|_1\ . \tag{A.10a}$$

$$C'_{\sigma,T} = \left\| \sigma^{1/2} \left(T - \mathrm{tr}(T\sigma) \mathbb{1} \right) \sigma^{1/2} \right\|_1\ , \tag{A.10b}$$

Intuitively, both these quantities $C_{\sigma,T}$ pick up the average charge absolute value (where T is the charge and according to the state σ), up to a constant charge offset μ or $\mathrm{tr}(T\sigma)$. These quantities generalize ΔT, $s(T)$, and $s'(T)$ by weighing the operator T with a general input state σ. The connection of (A.10) to these quantities, and the robustness of (A.10) with respect to perturbations of T, will be discussed in separate lemmas below (Lemmas 5 and 6).

Lemma 4. *Let $(\mathcal{N} \circ \mathcal{E})_{L\to A'}$ be the combined encoding and noise channel with total output system(s) A', where both encoding and noise channels may be any completely positive, trace-preserving maps. Let $\widehat{\mathcal{N} \circ \mathcal{E}}_{L\to E'}$ be a complementary channel with combined output system(s) E'. (In the context of Theorem II, we set $A' = A \otimes C$ and $E' = E \otimes C' \otimes F$, but this lemma holds more generally.) Suppose that there exist observables T_L and $Z_{E'}$ on the input and environment systems respectively such that*

$$\widehat{\mathcal{N} \circ \mathcal{E}}^\dagger (Z_{E'}) = T_L\ . \tag{A.11}$$

Then, for any $|\phi\rangle_{LR}$, both $\epsilon_{\mathrm{worst}}(\mathcal{N} \circ \mathcal{E})$ and $\epsilon_{|\phi\rangle}(\mathcal{N} \circ \mathcal{E})$ are lower bounded by two independent bounds:

$$\epsilon_{\mathrm{worst}}(\mathcal{N} \circ \mathcal{E}) \geqslant \epsilon_{|\phi\rangle}(\mathcal{N} \circ \mathcal{E}) \geqslant \begin{cases} \dfrac{C_{\phi_L,T_L}}{\Delta Z_{E'}} & \text{(A.12a)} \\[2ex] \dfrac{C'_{\phi_L,T_L}}{2\,\Delta Z_{E'}} & \text{(A.12b)} \end{cases}$$

Finally, if $\widehat{\mathcal{N} \circ \mathcal{E}}(\cdot) = \sum q_\alpha \widehat{\mathcal{N}_\alpha \circ \mathcal{E}}(\cdot)$ for a probability distribution $\{q_\alpha\}$ and a set of noise channels $\{\mathcal{N}_\alpha\}$, then for any $|\phi\rangle_{LR}$, the same bounds apply to the average of the individual error parameters corresponding to each erasure event:

$$\sum q_\alpha \epsilon_{|\phi\rangle}(\mathcal{N}_\alpha \circ \mathcal{E}) \geqslant \begin{cases} \dfrac{C_{\phi_L,T_L}}{\Delta Z_{E'}} & \text{(A.13a)} \\[2ex] \dfrac{C'_{\phi_L,T_L}}{2\,\Delta Z_{E'}} & \text{(A.13b)} \end{cases}$$

Proof of Lemma 4. We start by showing the following two statements: For any $|\phi\rangle_{LR}$, and for any state $\zeta_{E'}$, it holds that

$$\delta\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}), \zeta_{E'}\otimes\phi_R\Big) \geqslant \frac{C_{\phi_L,T_L}}{\Delta Z_{E'}}\ ;\ \text{and} \tag{A.14}$$

$$\delta\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}), \rho_{E'}\otimes\phi_R\Big) \geqslant \frac{C'_{\phi_L,T_L}}{\Delta Z_{E'}}\ , \tag{A.15}$$

where $\rho_{E'} = \widehat{\mathcal{N}\circ\mathcal{E}}(\phi_L)$.

We recall the following expressions for the one-norm of any Hermitian operator A:

$$\|A\|_1 = \max_{\|X\|_\infty\leqslant 1} \operatorname{tr}(XA) \tag{A.16a}$$

$$= \min_{\substack{\Delta_\pm\geqslant 0\\ A=\Delta_+-\Delta_-}} \operatorname{tr}(\Delta_+)+\operatorname{tr}(\Delta_-)\ , \tag{A.16b}$$

where the first optimization ranges over Hermitian operators X, and the second over positive semidefinite operators $\Delta_\pm$. We start form the left-hand side of (A.14). We choose a candidate X in (A.16a) of the form $Q_{E'}\otimes X'_R$ with $\|Q_{E'}\|_\infty\leqslant 1$ and $\|X'_R\|_\infty\leqslant 1$ and where Q_E and X'_R are Hermitian operators to be determined later. Then for any $|\phi\rangle_{LR}$ and for any $\zeta_{E'}$, we have

$$\begin{aligned}
&\frac{1}{2}\big\|\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}) - \zeta_{E'}\otimes\phi_R\big\|_1\\
&\qquad\geqslant \max_{\|X'_R\|_\infty\leqslant 1} \frac{1}{2}\operatorname{tr}\Big\{(Q_{E'}\otimes X'_R)\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}) - \zeta_{E'}\otimes\phi_R\Big)\Big\}\\
&\qquad= \max_{\|X'_R\|_\infty\leqslant 1} \frac{1}{2}\operatorname{tr}\Big\{\big(\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(Q_{E'})\otimes X'_R\big)\,\phi_{LR} - (Q_{E'}\zeta_{E'})\otimes(X'_R\,\phi_R)\Big\}\ ,
\end{aligned} \tag{A.17}$$

where the optimization ranges over Hermitian operators X'_R on the R system. Now we worry about choosing $Q_{E'}$. From Proposition 25, there exists $z\in\mathbb{R}$ such that $\Delta Z_{E'} = 2\|Z_{E'} - z\mathbb{1}\|_\infty$. We may assume without loss of generality that $z=0$, because we can shift $Z_{E'}\to Z_{E'}+z\mathbb{1}$ and $T_L\to T_L+z\mathbb{1}$ without any consequences for the assumptions or claims of the Lemma (noting that the quantities (A.10) are invariant under shifts of T by a multiple of the identity). Then we have $\Delta Z_{E'} = 2\|Z_{E'}\|$. Then define

$$Q_{E'} = \frac{Z_{E'}}{\|Z_{E'}\|_\infty} = \frac{2}{\Delta Z_{E'}}\,Z_{E'}\ . \tag{A.18}$$

With the main assumption of this lemma we have $\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(Q_{E'}) = (2/\Delta Z_{E'})\,T_L$. If we further restrict the optimization in (A.17) to operators X'_R such that $\operatorname{tr}(X'_R\phi_R)=0$, we obtain

$$\text{(A.17)} \geqslant \frac{1}{\Delta Z_{E'}} \max_{\substack{\|X'_R\|_\infty\leqslant 1\\ \operatorname{tr}(X'_R\phi_R)=0}} \operatorname{tr}\big\{T_L\operatorname{tr}_R(X'_R\,\phi_{LR})\big\}\ . \tag{A.19}$$

Without loss of generality, we may assume that $R\simeq L$ (if R is smaller, then embed it trivially in a larger system of same dimension as L; if R is larger, then remove unused dimensions on which ϕ_R has no support, noting that the support of ϕ_R may not exceed the dimension of L). Let $\{|k\rangle_L\}$, $\{|k\rangle_R\}$ be Schmidt bases of L and R corresponding to $|\phi\rangle_{LR}$, and recall that we have the relations $|\phi\rangle_{LR} = \phi_L^{1/2}\,|\Phi\rangle_{L:R} = \phi_R^{1/2}\,|\Phi\rangle_{L:R}$, where $|\Phi\rangle_{L:R} = \sum|k\rangle_L\otimes|k\rangle_R$ and where as before $\phi_L = \mathrm{tr}_R(\phi_{LR})$ and $\phi_R = \mathrm{tr}_L(\phi_{LR})$. Note that for any operator X'_R, we have $X'_R|\Phi\rangle_{L:R} = X_L|\Phi\rangle_{L:R}$ where X_L is related to X'_R by a transpose with respect to the bases used to define $|\Phi\rangle_{L:R}$, which implies also $\|X_L\|_\infty = \|X_R\|_\infty$. Consequently, $\mathrm{tr}_R(X'_R\phi_{LR}) = \mathrm{tr}_R(X'_R\,\phi_L^{1/2}\,\Phi_{L:R}\,\phi_L^{1/2}) = \phi_L^{1/2}\,X_L\,\phi_L^{1/2}$. Finally, note that $\mathrm{tr}(X'_R\phi_R) = \mathrm{tr}(X'_R\phi_{LR}) = \mathrm{tr}(\phi_L^{1/2}\,X_L\,\phi_L^{1/2}) = \mathrm{tr}(X_L\phi_L)$. So we obtain

$$
\begin{aligned}
\text{(A.19)} &= \frac{1}{\Delta Z_{E'}} \max_{\substack{\|X_L\|_\infty\leqslant 1\\ \mathrm{tr}(X_L\phi_L)=0}} \mathrm{tr}\big(\phi_L^{1/2}\,T_L\,\phi_L^{1/2}\,X_L\big)\ . \\
&= \frac{1}{\Delta Z_{E'}} \min_{\mu\in\mathbb{R}} \big\|\phi_L^{1/2}(T_L-\mu\mathbb{1})\phi_L^{1/2}\big\|_1 = \frac{C_{\phi_L,T_L}}{\Delta Z_{E'}}\ ,
\end{aligned}
\tag{A.20}
$$

where we invoke Proposition 26 and recall (A.10a), thus proving (A.14).

Now we show (A.15). Picking up from (A.17) while setting $\zeta_{E'} = \rho_{E'}$, we have

$$
\begin{aligned}
&\frac{1}{2}\big\|\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}) - \rho_{E'}\otimes\phi_R\big\|_1 \\
&\qquad\geqslant \max_{\|X'_R\|_\infty\leqslant 1} \frac{1}{\Delta Z_{E'}}\,\mathrm{tr}\Big\{(Z_{E'}\otimes X'_R)\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}-\phi_L\otimes\phi_R)\Big)\Big\} \\
&\qquad= \max_{\|X'_R\|_\infty\leqslant 1} \frac{1}{\Delta Z_{E'}}\,\mathrm{tr}\Big\{\big(\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(Z_{E'})\otimes X'_R\big)\,\big(\phi_{LR}-\phi_L\otimes\phi_R\big)\Big\}\ .
\end{aligned}
\tag{A.21}
$$

Again, we have $\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(Z_{E'}) = T_L$, and using the same procedure to define $|\Phi\rangle_{L:R}$ as above with X_L in one-to-one correspondence with X'_R via the transpose operation and with $\mathrm{tr}(X_L\phi_L) = \mathrm{tr}(X_R\phi_R)$, we obtain

$$
\begin{aligned}
\text{(A.21)} &= \max_{\|X_L\|_\infty\leqslant 1} \frac{1}{\Delta Z_{E'}}\Big[\mathrm{tr}\big\{T_L\phi_L^{1/2}X_L\phi_L^{1/2}\big\} - \mathrm{tr}\big\{T_L\phi_L\big\}\ \mathrm{tr}(X_L\phi_L)\Big] \\
&= \max_{\|X_L\|_\infty\leqslant 1} \frac{1}{\Delta Z_{E'}}\,\mathrm{tr}\Big\{X_L\Big(\phi_L^{1/2}\,(T_L-\mathrm{tr}(T_L\phi_L)\mathbb{1})\,\phi_L^{1/2}\Big)\Big\} \\
&= \frac{1}{\Delta Z_{E'}}\Big\|\phi_L^{1/2}\,\big(T_L-\mathrm{tr}(T_L\phi_L)\mathbb{1}_L\big)\,\phi_L^{1/2}\Big\|_1\ .
\end{aligned}
\tag{A.22}
$$

This proves (A.15).

Now, following Bény and Oreshkov [30], we have the duality of the entanglement fidelity also for a fixed input state, and there exists a state $\zeta_{E'}$ such that[1]

$$
F_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E},\mathrm{id}) = F(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta_{E'}\otimes\phi_R)\ ,
\tag{A.23}
$$

and thus

$$
\epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) = P(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta_{E'}\otimes\phi_R)\ ,
\tag{A.24}
$$

where $P(\sigma,\rho) = \sqrt{1-F^2(\sigma,\rho)}$ denotes the "purified distance" or "root infidelity" between the two states [35, 104, 105]. We recall the following known inequalities between this distance

[1] The statement with fixed input state is only briefly stated towards the end of their paper, as that claim is in fact easier to prove than their main theorem for the worst-case entanglement fidelity.

measure and the trace distance (see e.g. [104]):

$$\delta(\rho,\rho') \leqslant P(\rho,\rho') \leqslant \sqrt{2\delta(\rho,\rho')} \qquad \text{for all quantum states } \rho,\rho'. \tag{A.25}$$

We then have

$$\epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) = P\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta_{E'}\otimes\phi_R\big) \geqslant \delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta_{E'}\otimes\phi_R\big)\ , \tag{A.26}$$

which in combination with (A.14) proves (A.12a). The lower bound on ϵ_{worst} follows trivially from the fact that $\epsilon_{\text{worst}}(\cdot) = \max_{|\phi\rangle}\epsilon_{|\phi\rangle}(\cdot)$.

From (A.24), and using the fact that the purified distance cannot increase under partial trace, we find with $\rho_{E'} = \widehat{\mathcal{N}\circ\mathcal{E}}(\phi_L)$,

$$P(\rho_{E'},\zeta_{E'}) \leqslant \epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E})\ . \tag{A.27}$$

By triangle inequality, and using again (A.25),

$$\begin{aligned}\delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\rho_{E'}\otimes\phi_R\big) &\leqslant P\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\rho_{E'}\otimes\phi_R\big)\\ &\leqslant P\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta_{E'}\otimes\phi_R\big) + P\big(\zeta_{E'}\otimes\phi_R,\rho_{E'}\otimes\phi_R\big)\\ &\leqslant 2\,\epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E})\ .\end{aligned} \tag{A.28}$$

Combining this with (A.15) proves (A.12b).

Now we further assume that $\widehat{\mathcal{N}\circ\mathcal{E}} = \sum q_\alpha \widehat{\mathcal{N}_\alpha\circ\mathcal{E}}$ for some set of α's and a probability distribution $\{q_\alpha\}$. Then as above, invoking Bény and Oreshkov for each α with corresponding optimal states $\zeta^\alpha_{E'}$, we have

$$\begin{aligned}\sum q_\alpha\epsilon_\phi(\mathcal{N}_\alpha\circ\mathcal{E}) &= \sum q_\alpha P\big(\widehat{\mathcal{N}_\alpha\circ\mathcal{E}}(\phi_{LR}),\zeta^\alpha_{E'}\otimes\phi_R\big) \geqslant \sum q_\alpha \delta\big(\widehat{\mathcal{N}_\alpha\circ\mathcal{E}}(\phi_{LR}),\zeta^\alpha_{E'}\otimes\phi_R\big)\\ &\geqslant \delta\Big(\sum q_\alpha\widehat{\mathcal{N}_\alpha\circ\mathcal{E}}(\phi_{LR}),\sum q_\alpha\zeta^\alpha_{E'}\otimes\phi_R\Big) = \delta\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\zeta'_{E'}\otimes\phi_R\Big)\ ,\end{aligned} \tag{A.29}$$

using the joint convexity of the trace distance and defining $\zeta'_{E'} = \sum q_\alpha\zeta^\alpha_{E'}$. Directly invoking (A.14) then proves (A.13a). We also have (A.29) $\geqslant \delta(\rho_{E'},\zeta'_{E'})$, and hence by triangle inequality

$$\delta\Big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}),\rho_{E'}\otimes\phi_R\Big) \leqslant 2\sum q_\alpha\epsilon_\phi(\mathcal{N}_\alpha\circ\mathcal{E})\ . \tag{A.30}$$

Combining with (A.15) then yields (A.13b). ■

First, we consider special cases for the bounding quantities (A.10) and relate them to ΔT_L, $s(T_L)$, and $s'(T_L)$:

Lemma 5. *Let $|\psi^\pm\rangle$ be eigenstates associated with the maximum and the minimum eigenvalues of T_L, respectively, and let $\Pi_L = |\psi^+\rangle\langle\psi^+| - |\psi^-\rangle\langle\psi^-|$. The quantities (A.10) take* *the following values for selected choices of input states:*

$$C_{\mathbb{1}/d_L,T_L} = s(T_L) \ ; \tag{A.31a}$$
$$C'_{\mathbb{1}/d_L,T_L} = s'(T_L) \ ; \tag{A.31b}$$
$$C_{\Pi_L/2,T_L} = \Delta T_L/2 \ . \tag{A.31c}$$

Proof of Lemma 5. Proposition 27 immediately yields the first equality. The second equality also follows immediately by plugging in $\sigma = \mathbb{1}_L/d_L$ into (A.10b). For the third equality, we first invoke Proposition 26 to write

$$C_{\Pi_L/2,T_L} = \frac{1}{2} \max_{\substack{\|X_L\|_\infty \leqslant 1 \\ \mathrm{tr}(X_L\Pi_L)=0}} \mathrm{tr}(\Pi_L T_L \Pi_L X_L) \ . \tag{A.32}$$

We choose as candidate $X_L = |\psi^+\rangle\langle\psi^+|_L - |\psi^-\rangle\langle\psi^-|_L$, since we have indeed $\mathrm{tr}(\Pi_L X_L) = 0$ and $\|X_L\|_\infty \leqslant 1$, and we obtain

$$C_{\Pi_L/2,T_L} \geqslant \frac{1}{2}\,\mathrm{tr}\big(\Pi_L T_L \Pi_L X_L\big) = \frac{\Delta T_L}{2} \ . \tag{A.33}$$

To complete the proof we also need to show the opposite bound. Let $\nu = \big[\langle\psi^+|T_L|\psi^+\rangle + \langle\psi^-|T_L|\psi^-\rangle\big]/2$ be the average between the maximum and the minimum eigenvalue of T_L. From Proposition 26,

$$C_{\Pi_L/2,T_L} \leqslant \frac{1}{2}\big\|\Pi_L\,(T_L - \nu\mathbb{1})\,\Pi_L\big\|_1 = \frac{\Delta T_L}{2} \ , \tag{A.34}$$

noting that the one-norm with the projector Π_L picks out the sum of the gaps between the two extremal eigenvalues of T_L and their midpoint. ∎

We may now combine the above lemmas to finally prove Theorem II.

Proof of Theorem II. Thanks to Lemma 3 there exists $Z_{C'E}$ such that

$$\widehat{\mathcal{N}\circ\mathcal{E}}^\dagger(Z_{C'E}) = T_L \ ; \qquad\qquad \Delta Z_{C'E} \leqslant \max_{\alpha\in K'} \frac{\Delta T_\alpha}{q_\alpha} \ . \tag{A.35}$$

We may directly plug this observable into Lemma 4 to deduce that the bound (A.12a) applies to our approximately covariant code. We now express this bound for the particular quantities $\epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E})$, $\langle\epsilon_{\mathrm{e}}(\mathcal{N}^\alpha\circ\mathcal{E})\rangle_\alpha$ and $\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E})$.

First, let $|\phi\rangle_{LR} = |\hat{\phi}\rangle_{LR}$ be the maximally entangled state between L and $R \simeq L$. Then Lemma 4 states that both quantities $C_{\mathbb{1}/d_L,T_L}/\Delta Z_{C'E}$ and $[C'_{\mathbb{1}/d_L,T_L}/2]/\Delta Z_{C'E}$ are lower bounds to $\epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E})$ and to $\langle\epsilon_{\mathrm{e}}(\mathcal{N}^\alpha\circ\mathcal{E})\rangle_\alpha$, which proves (A.3a) as we recall the property (A.35) and we use Lemma 5.

For $\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E})$, we get to pick $|\phi\rangle_{LR}$ freely and this will yield a valid bound. Let $|\psi^\pm\rangle_L$ be eigenstates of T_L corresponding to the maximum and minimum eigenvalues T_L, respectively, with $\langle\psi^+|T_L|\psi^+\rangle - \langle\psi^-|T_L|\psi^-\rangle = \Delta T_L$. Now choose two arbitrary orthogonal states $|\pm\rangle_R$

on R and set

$$|\phi\rangle_{LR} = \frac{1}{\sqrt{2}}\big[|\psi^+\rangle_L|+\rangle_R + |\psi^-\rangle_L|-\rangle_R\big] \ . \tag{A.36}$$

Lemmas 4 and 5 then assert that

$$\epsilon_{\text{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) \geqslant \frac{C_{\Pi_L/2,T_L}}{\Delta Z_{C'E}} \geqslant \frac{\Delta T_L/2}{\max_{\alpha\in K'}\Delta T_\alpha/q_\alpha}\ , \tag{A.37}$$

where $\Pi_L = |\psi^+\rangle\langle\psi^+|_L + |\psi^-\rangle\langle\psi^-|_L$. This proves (A.3b). ■

Lemma 6. *The quantities (A.10) are robust to perturbations of T. Namely, let σ be any state, let $\delta' \geqslant 0$ and let T, T' be two Hermitian operators such that $\|T - T'\|_\infty \leqslant \delta'$. Then*

$$\big|C_{\sigma,T'} - C_{\sigma,T}\big| \leqslant \delta'\ ; \qquad\qquad \big|C'_{\sigma,T'} - C'_{\sigma,T}\big| \leqslant 2\delta'\ . \tag{A.38}$$

The quantities are also somewhat robust under perturbations of σ, as can be seen from the expression (G.3b) in Proposition 26, but the corresponding error terms for $C_{\sigma,T}$ will necessarily have to depend on features of T (no uniform continuity). For instance, consider on a qubit the operator $T = c|0\rangle\langle 0|$ for some large c and $\sigma = |1\rangle\langle 1|$, $\sigma' = \varepsilon|0\rangle\langle 0| + (1-\varepsilon)|1\rangle\langle 1|$; then $C_{\sigma,T} = 0$ and $C_{\sigma',T} = c\epsilon$ as can be seen by choosing $\mu = 0$ and $X = |0\rangle\langle 0| - [\varepsilon/(1-\varepsilon)]|1\rangle\langle 1|$ in Proposition 26. While $C_{\sigma,T}$ is still differentiable with respect to this shift in σ, the derivative can be made arbitrarily large by a suitable choice of c.

Proof of Lemma 6. For any $t \in \mathbb{R}$, we have

$$\|\sigma^{1/2}\,(T' - t\mathbb{1})\,\sigma^{1/2}\|_1 \leqslant \|\sigma^{1/2}\,(T - t\mathbb{1})\,\sigma^{1/2}\|_1 + \|\sigma^{1/2}\,(T - T')\,\sigma^{1/2}\|_1\ . \tag{A.39}$$

The second term on the right hand side can be bounded as follows. Let $G_\pm \geqslant 0$ be the positive and negative parts of $T - T'$ such that $T - T' = G_+ - G_-$, noting that $\|G_\pm\|_\infty \leqslant \delta'$ because all eigenvalues of $T - T'$ must be less than δ' in magnitude. Then define $\Delta_\pm = \sigma^{1/2}\,G_\pm\,\sigma^{1/2} \geqslant 0$, such that $\Delta_+ - \Delta_- = \sigma^{1/2}(T - T')\sigma^{1/2}$. Recall the expression (A.16b) for the one-norm of a Hermitian matrix. Then

$$\|\sigma^{1/2}\,(T - T')\,\sigma^{1/2}\|_1 \leqslant \operatorname{tr}(\Delta_+) + \operatorname{tr}(\Delta_-) \leqslant \operatorname{tr}\big(\sigma(G_+ + G_-)\big) \leqslant \|G_+ + G_-\|_\infty \leqslant \delta'\ , \tag{A.40}$$

since G_+ and G_- have non-overlapping support and $\|G_\pm\|_\infty \leqslant \delta'$. Choosing t to be optimal for $C_{\sigma,T}$, Eq. (A.39) becomes

$$C_{\sigma,T'} \leqslant C_{\sigma,T} + \delta'\ . \tag{A.41}$$

We may interchange the roles of T and T', and this shows the first part of claim. For the quantity (A.10a), we further note that $\operatorname{tr}\big((T - T')\sigma\big) \leqslant \big\|\sigma^{1/2}\big(T - T'\big)\sigma^{1/2}\big\|_1 \leqslant \delta'$ and therefore

$$\begin{aligned}\|\sigma^{1/2}\,(T' - \operatorname{tr}(T'\sigma)\mathbb{1})\,\sigma^{1/2}\|_1 &\leqslant \|\sigma^{1/2}\,(T - \operatorname{tr}(T\sigma)\mathbb{1})\,\sigma^{1/2}\|_1 + \|\sigma^{1/2}\,(T' - T)\,\sigma^{1/2}\|_1 \\ &\quad + \big|\operatorname{tr}(T'\sigma) - \operatorname{tr}(T\sigma)\big|\,\|\sigma\|_1 \\ &\leqslant \|\sigma^{1/2}\,(T - \operatorname{tr}(T\sigma)\mathbb{1})\,\sigma^{1/2}\|_1 + 2\delta'\ .\end{aligned} \tag{A.42}$$

Finishing the argument as for $C_{\sigma,T}$ proves the robustness claim of $C'_{\sigma,T}$. ■

We may now give a proof of Corollary III.

Proof of Corollary III. The charge T_A is $(\{t_\alpha^\pm\}_{\alpha\in K'},\eta)$-bounded for $\mathcal{E}$. Let $\Pi_\alpha = \mathbb{1} - \Pi_\alpha^\perp$ be the projector which projects onto the eigenspaces of T_α whose corresponding eigenvalues are in the range $[t_\alpha^-, t_\alpha^+]$. We may assume without loss of generality that $t_\alpha^- = -t_\alpha^+$ by shifting $T_\alpha \to T_\alpha + c\mathbb{1}$ and $t_\alpha^\pm \to t_\alpha^\pm + c$ as required without impacting our assumptions or claims. Then $\Delta t_\alpha = t_\alpha^+ - t_\alpha^- = 2t_\alpha^+$. Define the observables $\tilde{T}_\alpha = \Pi_\alpha T_\alpha$, and observe that $\Delta\tilde{T}_\alpha \leqslant \Delta t_\alpha$, since $\tilde{T}_\alpha$ has eigenvalues between $-\Delta t_\alpha/2$ and $+\Delta t_\alpha/2$.

Let $\tilde{T}_A = \sum_\alpha \tilde{T}_\alpha$, and define $\Xi_A = T_A - \tilde{T}_A = \sum_\alpha(\mathbb{1}-\Pi_\alpha)T_\alpha = \sum_\alpha \Pi_\alpha^\perp T_\alpha$. For any logical state σ_L,

$$\left|\mathrm{tr}\big(\mathcal{E}(\sigma_L)\,\Xi_A\big)\right| = \left|\mathrm{tr}\Big(\mathcal{E}(\sigma_L)\sum \Pi_\alpha^\perp T_\alpha\Big)\right| \leqslant \eta\ , \tag{A.43}$$

using the assumption that T_A is $(\{t_\alpha^\pm\}_{\alpha\in K'},\eta)$-bounded for $\mathcal{E}$. Therefore $\|\mathcal{E}^\dagger(\Xi_A)\|_\infty \leqslant \eta$, as we can choose σ_L to be a state supported on the eigenspace associated with the extremal eigenvalue of $\mathcal{E}^\dagger(\Xi_A)$ thus satisfying $|\mathrm{tr}(\sigma_L\mathcal{E}^\dagger(\Xi_A))| = \|\mathcal{E}^\dagger(\Xi_A)\|_\infty$. Now let

$$\tilde{T}_L = \mathcal{E}^\dagger\big(\tilde{T}_A\big)\ . \tag{A.44}$$

Then, since $\mathcal{E}$ is δ-charge-conserving, we may assume without loss of generality that $\left\|\mathcal{E}^\dagger(T_A) - T_L\right\|_\infty \leqslant \delta$, because we may shift T_L by a multiple of the identity without impacting our assumptions or claims. Then

$$\|\tilde{T}_L - T_L\|_\infty = \|\mathcal{E}^\dagger\big(\Xi_A\big) + \mathcal{E}^\dagger\big(T_A\big)\|_\infty \leqslant \delta + \eta\ . \tag{A.45}$$

We now invoke Lemma 3 for $\tilde{T}_L$ instead of T_L and $\tilde{T}_A = \sum \tilde{T}_\alpha$ instead of T_A. Combining with Lemma 4 yields, for any $|\phi\rangle_{LR}$,

$$\begin{aligned}
\left.\begin{array}{r}\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \epsilon_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) \\[2ex] \sum q_\alpha \epsilon_{|\phi\rangle}\big(\mathcal{N}_\alpha\circ\mathcal{E}\big)\end{array}\right\} &\geqslant \frac{\max\big\{C_{\phi_L,\tilde{T}_L}, C'_{\phi_L,\tilde{T}_L}/2\big\}}{\max_{\alpha\in K'}\ q_\alpha^{-1}\Delta\tilde{T}_\alpha} \\
&\geqslant \frac{\max\big\{C_{\phi_L,T_L}, C'_{\phi_L,T_L}/2\big\} - \delta - \eta}{\max_{\alpha\in K'}\ q_\alpha^{-1}\Delta t_\alpha}\ ,
\end{aligned} \tag{A.46}$$

where the last inequality holds because $C_{\sigma,T}, C'_{\sigma,T}$ are robust to perturbations of T (Lemma 6), and because $\Delta\tilde{T}_\alpha = \Delta t_\alpha$.

Using Lemma 5 and taking the same special cases as in the proof of Theorem II proves the claim. ■

At this point we comment on Condition (A.2) in the assumptions of Corollary III. It may look a bit awkward, but its meaning is intuitively simple: First, we need to shift the charge values to center them at zero for each α for our proof. Second, we need to make sure that if we project any codeword into the given range of physical charge values for each α, then the total error we make when attempting to determine the expectation value of the actual (possibly unbounded) charge observable T_A is small. In practice, this just means that the part of the codewords outside of the given range of charge values only has a small contribution to the total expectation value of charge. For convenience we may use the following simplified criterion, where we simply fix a charge cut-off value t:

Proposition 7. *Consider $\mathcal{E}_{L\to A}$, T_L, K' and $T_A = \sum_{\alpha\in K'} T_\alpha$ as in Corollary III. Let $t > 0$. Set $t_\alpha^+ = -t_\alpha^- = t$ and define $\Pi_\alpha, \Pi_\alpha^\perp$ as above. Let $\{|\phi_\alpha^{t',j}\rangle\}$ be an eigenbasis of T_α corresponding to eigenvalues t' with a possible degeneracy index j. Suppose that there is an $\eta' \geqslant 0$ such that for any logical state ψ_L and for any α,*

$$\sum_{t',j:\,|t'|>t} |t'|\,\langle\phi_\alpha^{t',j}\,|\,\rho_\alpha\,|\,\phi_\alpha^{t',j}\rangle \leqslant \eta' \ , \tag{A.47}$$

where we write $\rho_\alpha = \mathrm{tr}_{A\setminus A_\alpha}(\mathcal{E}(\psi_L))$ and where the sum ranges over the eigenstate labels (t', j) such that $|t'| > t$. Then, condition (A.2) is satisfied with $\eta = |K'|\,\eta'$, and furthermore $\Delta T_\alpha = 2t$ for all α.

Proof of Proposition 7. We have $t_\alpha = (t_\alpha^+ + t_\alpha^-)/2 = 0$. For any ψ_L, calculate

$$\begin{aligned}
\Big|\sum \mathrm{tr}\big(\Pi_\alpha^\perp T_\alpha\,\mathcal{E}(\psi_L)\big)\Big| &\leqslant \sum \big|\mathrm{tr}\big(\Pi_\alpha^\perp T_\alpha\,\mathcal{E}(\psi_L)\big)\big| \\
&\leqslant \sum_\alpha \Big|\sum\nolimits_{t',j:\,|t'|>t} t'\,\langle\phi_\alpha^{t',j}\,|\,\mathrm{tr}_{A\setminus A_\alpha}(\mathcal{E}(\psi_L))\,|\,\phi_\alpha^{t',j}\rangle\Big| \\
&\leqslant \sum_\alpha \sum\nolimits_{t',j:\,|t'|>t} |t'|\,\langle\phi_\alpha^{t',j}\,|\,\mathrm{tr}_{A\setminus A_\alpha}(\mathcal{E}(\psi_L))\,|\,\phi_\alpha^{t',j}\rangle \\
&\leqslant \sum_\alpha \eta' \leqslant |K'|\,\eta' \ .
\end{aligned} \tag{A.48}$$

Note by the way that the left hand side of (A.47) is exactly $\mathrm{tr}(\Pi_\alpha^\perp\,|T_\alpha|\mathcal{E}(\psi_L))$. ■

SM. B: Correlation functions and bounds

In this section we present an alternative strategy for proving the bound (27), by studying the connected correlation functions between the physical subsystems and the logical information.

The covariance of the codes can be seen as a linear constraint, which can be easily employed to obtain a second order constraints. To start, we again assume the simpler case of isometric encoding. We construct the state corresponding to the encoding isometry $V_{L\to A}$ by injecting a maximally entangled state $|\hat{\phi}\rangle_{LR}$ to $V_{L\to A}$ (Fig. 1):

$$|\Psi\rangle_{AR} = V|\hat{\phi}\rangle_{LR} \ . \tag{B.1}$$

We have $T_A|\Psi\rangle_{LA} = T_A V|\hat{\phi}\rangle_{LR} = V(T_L - \nu\mathbb{1}_L)|\hat{\phi}\rangle_{LR}$ for some constant ν. Define $T_R = (T_L - \nu\mathbb{1}_L)^T$ where the transpose is taken as a matrix ignoring the Hilbert space label; this ensures that $(T_L - \nu\mathbb{1}_L)|\hat{\phi}\rangle_{LR} = T_R|\hat{\phi}\rangle_{LR}$. Therefore, the covariance of V translates to the invariance of $|\Psi\rangle$:

$$\left(\sum_{i=1}^n T_{A_i}\right)|\Psi\rangle_{RA} = T_A|\Psi\rangle_{RA} = T_R|\Psi\rangle_{RA} \ . \tag{B.2}$$

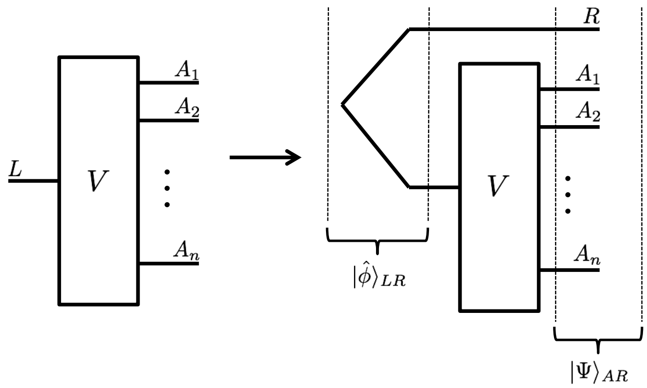

FIG. 1: Depiction of the construction of the state $|\Psi\rangle_{AR}$ by injecting the maximally entangled state $|\hat{\phi}\rangle_{LR}$ into the encoding isometry $V_{L\to A}$.

We define the *connected correlator* between two operators A, B as

$$\langle A, B\rangle := \mathrm{tr}\,(AB\Psi) - \mathrm{tr}\,(A\Psi)\,\mathrm{tr}\,(B\Psi)\ . \tag{B.3}$$

Consider an arbitrary operator X_R. It be seen from (B.2) that

$$\langle X_R, T_R\rangle = \sum_{i=1}^{n} \langle X_R, T_{A_i}\rangle\ .$$

Using the triangle inequality, we obtain

$$|\langle X_R, T_R\rangle| \leqslant \sum_{i=1}^{n} |\langle X_R, T_{A_i}\rangle| \qquad \text{for all } X_R. \tag{B.4}$$

Although the derivation of (B.4) is very simple, it provides a general lower bound to the amount of correlations between the reference system and the physical subsystems, from which we can draw physical consequences. The correlation functions measure how close the state Ψ_{RA_i} is to the product state $\Psi_R \otimes \Psi_{A_i}$:

$$\begin{aligned} |\langle X_R, T_{A_i}\rangle| &= |\mathrm{tr}[X_R T_{A_i}(\Psi_{RA_i} - \Psi_R \otimes \Psi_{A_i})]| \\ &\leqslant \|X_R\|_\infty \, \|T_{A_i}\|_\infty \, \|\Psi_{RA_i} - \Psi_R \otimes \Psi_{A_i}\|_1\ , \end{aligned} \tag{B.5}$$

where we used Hölder's inequality. We can replace $T_{A_i} \to T_{A_i} - t\mathbb{1}$ in (B.5) without changing the left hand side of the inequality as $\langle X_R, T_{A_i} - t\mathbb{1}\rangle = \langle X_R, T_{A_i}\rangle$:

$$\begin{aligned} |\langle X_R, T_{A_i}\rangle| &\leqslant \|X_R\|_\infty \, \|T_{A_i} - t\mathbb{1}\|_\infty \, \|\Psi_{RA_i} - \Psi_R \otimes \Psi_{A_i}\|_1 \\ &= \frac{1}{2}\|X_R\|_\infty \, \Delta T_{A_i} \, \|\Psi_{RA_i} - \Psi_R \otimes \Psi_{A_i}\|_1\ , \end{aligned} \tag{B.6}$$

where the second line follows by a suitable choice of t, and where ΔT_{A_i} is the difference between the maximum and minimum eigenvalue of T_{A_i}.

The accuracy to which the code V can correct against errors is precisely determined by how close Ψ_{RA_i} is to a product state. Indeed, consider the noise channel $\mathcal{N}^i_{A\to A}$ in (A.1) that erases the system A_i. By Bény and Oreshkov (12a), we have

$$\begin{aligned}
\epsilon_{\mathrm{e}}(\mathcal{N}^i\circ\mathcal{E}) &= \min_\zeta \sqrt{1-F^2\big(\widehat{\mathcal{N}^i\circ\mathcal{E}}(\hat\phi_{LR}),\zeta\otimes\Psi_R\big)} \\
&\geqslant \min_\zeta \frac{1}{2}\big\|\widehat{\mathcal{N}^i\circ\mathcal{E}}(\hat\phi_{LR})-\zeta\otimes\Psi_R\big\|_1 \\
&= \min_\zeta \frac{1}{2}\big\|\Psi_{RA_i}-\zeta_{A_i}\otimes\Psi_R\big\|_1
\end{aligned} \tag{B.7}$$

where $\widehat{\mathcal{N}^i\circ\mathcal{E}}(\hat\phi_{LR})=\Psi_{RA_i}$ and $\Psi_R=\mathbb{1}_R/d_L$, and where we have used the known relation $\delta(\cdot,\cdot)\leqslant\sqrt{1-F^2(\cdot,\cdot)}$ between the trace distance and the fidelity. Because the trace distance cannot increase under the partial trace, and if we set ζ_{A_i} to be the optimal state in the expression above, also have $(1/2)\|\Psi_{A_i}-\zeta_{A_i}\|_1\leqslant\epsilon_{\mathrm{e}}(\mathcal{N}^i\circ\mathcal{E})$ and thus by triangle inequality,

$$\frac{1}{2}\big\|\Psi_{RA_i}-\Psi_{A_i}\otimes\Psi_R\big\|_1\leqslant 2\epsilon_{\mathrm{e}}(\mathcal{N}^i\circ\mathcal{E})\ . \tag{B.8}$$

It remains to combine (B.8) with (B.5) and (B.4) and to choose the best possible X_R to get our final result.

Theorem 8. *The individual entanglement fidelities of recovery of a covariant code $\mathcal{E}(\cdot)=V(\cdot)V^\dagger$ against single erasures at known locations satisfy the following inequality:*

$$\frac{1}{2d_L}\left\|T_L-\mathrm{tr}(T_L)\frac{\mathbb{1}}{d_L}\right\|_1\leqslant\sum_{i=1}^n\Delta T_i\,\epsilon_{\mathrm{e}}(\mathcal{N}^i\circ\mathcal{E})\ . \tag{B.9}$$

Furthermore, this can be used to show that

$$\epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E})\geqslant\frac{1}{2d_L}\frac{\|T_L-\mathrm{tr}(T_L)\mathbb{1}/d_L\|_1}{\max_i q_i^{-1}\Delta T_i}\ . \tag{B.10}$$

Note that $T_L-\mathrm{tr}(T_L)\mathbb{1}/d_L$ is just a shift of T_L by a multiple of identity to make it traceless. Therefore, $\|T_L-\mathrm{tr}(T_L)\mathbb{1}/d_L\|_1$ is a 1-norm measure for the spread of eigenvalues of T_L. The bounds of Theorem 8 and Eq. (27) have a very similar nature.

Proof of Theorem 8. We start with the correlator in the left hand side of (B.4):

$$\langle X_R,T_R\rangle=\mathrm{tr}(X_RT_R\Psi_R)-\mathrm{tr}(X_R\Psi_R)\,\mathrm{tr}(T_R\Psi_R)=\frac{1}{d_L}\,\mathrm{tr}\bigg(X_R\bigg[T_R-\frac{\mathrm{tr}(T_R)}{d_L}\mathbb{1}\bigg]\bigg)\ . \tag{B.11}$$

Now, choose the optimal X_R such that $\|X_R\|_\infty\leqslant 1$ and that $\|T_R-\mathrm{tr}(T_R)\mathbb{1}/d_L\|_1=\mathrm{tr}[X_R(T_R-\mathrm{tr}(T_R)\mathbb{1}/d_L)]$. Plugging into (B.4), and combining with (B.6) and (B.8), immediately gives (B.9).

Furthermore from (B.9) we have

$$\frac{1}{d_L}\left\|T_L-\mathrm{tr}(T_L)\frac{\mathbb{1}}{d_L}\right\|_1\leqslant\sum(q_i^{-1}\Delta T_i)(q_i\epsilon_{\mathrm{e}}(\mathcal{N}^i\circ\mathcal{E}))$$

$$\leqslant \Big(\max_i \,(q_i^{-1}\Delta T_i)\Big) \sum q_i \epsilon_{\mathrm{e}}(\mathcal{N}^i \circ \mathcal{E}) \ . \tag{B.12}$$

By convexity of $x \mapsto x^2$, and by Lemma 29, we have

$$\sum q_i \epsilon_{\mathrm{e}}(\mathcal{N}^i \circ \mathcal{E}) \leqslant \sqrt{\sum q_i \epsilon_{\mathrm{e}}^2(\mathcal{N}^i \circ \mathcal{E})} = \epsilon_{\mathrm{e}}(\mathcal{N} \circ \mathcal{E}) \ . \tag{B.13}$$

Combining (B.13) with (B.12) proves (B.10). ∎

SM. C: Criterion for approximate codes

When we come up with a new code, how can we show that it forms an ϵ-approximate error-correcting code against erasures at known locations? Here we provide a criterion that, when it can be applied, certifies that a given code performs well.

Let L be the logical space and A be the physical space, and consider an encoding operation $\mathcal{E}_{L\to A}$ that can be any completely positive, trace-preserving map. Note that in the case of a more general noise model, A does not necessarily have to be composed of several subsystems. Consider a collection of noise channels $\{\mathcal{N}^\alpha\}$ and probabilities $\{q_\alpha\}$. We assume that the environment applies a random noise channel from this set with the corresponding probability, while providing a record of which noise channel was applied in a separate register C. The overall noise channel that is applied by the environment is then

$$\mathcal{N}_{A\to AC}(\cdot) = \sum q_\alpha |\alpha\rangle\langle\alpha|_C \otimes \mathcal{N}^\alpha_{A\to A}(\cdot) \ . \tag{C.1}$$

Given complementary channels $\widehat{\mathcal{N}^\alpha \circ \mathcal{E}}$ of $\mathcal{N}^\alpha \circ \mathcal{E}$, we can construct a complementary channel of $\mathcal{N} \circ \mathcal{E}$ as

$$\widehat{\mathcal{N} \circ \mathcal{E}}_{A\to C'E}(\cdot) = \sum q_\alpha |\alpha\rangle\langle\alpha|_{C'} \otimes \widehat{\mathcal{N}^\alpha \circ \mathcal{E}}(\cdot) \ , \tag{C.2}$$

with an additional register C' and where the outputs of the individual complementary channels for each α are embedded into a system E.

We fix any basis $\{|x\rangle_L\}$ of L, and we define for each α the operators

$$\rho_\alpha^{x,x'} = \widehat{\mathcal{N}^\alpha \circ \mathcal{E}}(|x\rangle\langle x'|_L) \ . \tag{C.3}$$

Note that $\rho_\alpha^{x,x}$ is a quantum state for each α and for each x, but that $\rho_\alpha^{x,x'}$ is not necessarily even Hermitian for $x \neq x'$.

For an isometric encoding $\mathcal{E}$, and in the noise $\mathcal{N}$ acts by erasing a collection of subsystems labeled by α and chosen with probability q_α, the operators $\rho_\alpha^{x,x'}$ are simply the reduced operators on the sites labeled by α of the logical operator $|x\rangle\langle x'|$:

$$\rho_\alpha^{x,x'} = \mathrm{tr}_{A\setminus A_\alpha}(\mathcal{E}(|x\rangle\langle x'|)) \ . \tag{C.4}$$

Proposition 9. *Assume that there exists $\nu, \epsilon' \geqslant 0$, and that there exists a quantum state ζ_α*

for each α, such that for all α,

$$F(\rho_\alpha^{x,x}, \zeta_\alpha) \geqslant \sqrt{1-\epsilon'^2} \quad \textit{for all } x \ ; \quad \textit{and} \tag{C.5a}$$

$$\|\rho_\alpha^{x,x'}\|_1 \leqslant \nu \qquad \textit{for all } x \neq x' \ . \tag{C.5b}$$

Then $\mathcal{E}_{L\to A}$ is an approximate error-correcting code against the noise $\mathcal{N}$, with approximation parameter

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \leqslant \epsilon' + d_L\sqrt{\nu} \ , \tag{C.6}$$

where d_L is the dimension of the logical system L.

Proof of Proposition 9. Let

$$\zeta_{C'E} = \sum_\alpha q_\alpha |\alpha\rangle\langle\alpha|_{C'} \otimes \zeta_\alpha \ . \tag{C.7}$$

Using the Bény-Oreshkov property (12), the proof strategy is to find a lower bound to the entanglement fidelity of the channel $\widehat{\mathcal{N}\circ\mathcal{E}}$ to the constant channel $\mathcal{T}_\zeta$ outputting the state $\zeta_{C'E}$ defined above.

Consider a reference system $R \simeq L$, and let $\{|x\rangle_R\}$ be any fixed basis of R. Let $|\Phi\rangle_{L:R} = \sum_x |x\rangle_L \otimes |x\rangle_R$. For any state $|\sigma\rangle_{LR}$, there exists a complex matrix B_R such that $|\sigma\rangle_{LR} = B_R\,|\Phi\rangle_{L:R}$ and $\sigma_R = \mathrm{tr}_L(\sigma_{LR}) = B_R B_R^\dagger$ (choose $B_R = \sum_{x,x'}\langle x, x'\,|\sigma\rangle_{LR} |x'\rangle\langle x|_R$). Note that $\|B_R B_R^\dagger\|_\infty = \|B_R^\dagger B_R\|_\infty \leqslant 1$. We have

$$\begin{aligned}(\widehat{\mathcal{N}\circ\mathcal{E}} \otimes \mathrm{id}_R)(\sigma_{LR}) &= B_R\,\widehat{\mathcal{N}\circ\mathcal{E}}(\Phi_{L:R})\,B_R^\dagger \\ &= \sum_\alpha q_\alpha |\alpha\rangle\langle\alpha|_{C'} \otimes (B_R\,\widehat{\mathcal{N}^\alpha\circ\mathcal{E}}(\Phi_{L:R})\,B_R^\dagger) \\ &= \sum_{\alpha,x,x'} q_\alpha\,|\alpha\rangle\langle\alpha|_{C'} \otimes \big(B_R\,\rho_{ER}^\alpha\,B_R^\dagger\big) \ ,\end{aligned} \tag{C.8}$$

where we have defined for each α the positive semidefinite operator

$$\rho_{ER}^\alpha = \widehat{\mathcal{N}^\alpha\circ\mathcal{E}}(\Phi_{L:R}) = \sum_{x,x'} \rho_\alpha^{x,x'} \otimes |x\rangle\langle x'|_R \ . \tag{C.9}$$

While the ρ_{ER}^α's are positive semidefinite, they are not normalized to unit trace as proper quantum states. Recalling that the fidelity is jointly concave, we have

$$\begin{aligned}F\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_{LR}), \zeta_{C'E}\otimes\sigma_R\big) &= F\Big(\textstyle\sum_\alpha q_\alpha|\alpha\rangle\langle\alpha|_{C'}\otimes(B_R\rho_{ER}^\alpha B_R^\dagger), \sum_\alpha q_\alpha|\alpha\rangle\langle\alpha|_{C'}\otimes\zeta_\alpha\otimes\sigma_R\Big) \\ &\geqslant \sum_\alpha q_\alpha\,F\big(B_R\rho_{BR}^\alpha B_R^\dagger, \zeta_\alpha\otimes\sigma_R\big) \ .\end{aligned} \tag{C.10}$$

At this point, we define for each α the positive semidefinite operator

$$\tilde\rho_{ER}^\alpha = \sum_x \rho_\alpha^{x,x} \otimes |x\rangle\langle x|_R \ . \tag{C.11}$$

Note that $B_R\tilde\rho_{ER}^\alpha B_R^\dagger$ is a quantum state, because $\mathrm{tr}(B_R\tilde\rho_{ER}^\alpha B_R^\dagger) = \sum_x \mathrm{tr}(B_R|x\rangle\langle x|B_R^\dagger) =$

$\mathrm{tr}(B_R B_R^\dagger) = 1$. In fact, the quantum states $B_R \tilde{\rho}^\alpha_{ER} B_R^\dagger$ and $B_R \rho^\alpha_{ER} B_R^\dagger$ are close in trace distance:

$$\begin{aligned}\left\| B_R \left(\rho^\alpha_{ER} - \tilde{\rho}^\alpha_{ER}\right) B_R^\dagger \right\|_1 &= \left\| B_R \left(\sum\nolimits_{x\neq x'} \rho_\alpha^{x,x'} \otimes |x\rangle\langle x'| \right) B_R^\dagger \right\|_1 \\ &= \left\| \sum\nolimits_{x\neq x'} \rho_\alpha^{x,x'} \otimes \left(B_R\, |x\rangle\langle x'|\, B_R^\dagger\right) \right\|_1 \\ &\leqslant \sum\nolimits_{x\neq x'} \left\| \rho_\alpha^{x,x'} \right\|_1 \cdot \| B_R\, |x\rangle\langle x'|\, B_R^\dagger \|_1 \\ &\leqslant \sum\nolimits_{x\neq x'} \left\| \rho_\alpha^{x,x'} \right\|_1 \leqslant d_L^2\, \nu \ , \end{aligned} \tag{C.12}$$

using our assumption (C.5b), and noting that $\|B_R\, |x\rangle\langle x'|\, B_R^\dagger\|_1 \leqslant \|B_R |x\rangle\|_1\, \|\langle x'| B_R^\dagger\|_1 = \|\langle x| B_R^\dagger\|_1\, \|\langle x'|B_R^\dagger\|_1 \leqslant 1$ because $\mathrm{tr}\sqrt{\langle x|\, B_R^\dagger B_R\, |x\rangle} \leqslant 1$. Recalling the relation $P(\cdot,\cdot) \leqslant \sqrt{2\delta(\cdot,\cdot)} = \sqrt{\|(\cdot) - (\cdot)\|_1}$ between the purified distance $P(\cdot,\cdot) = \sqrt{1 - F^2(\cdot,\cdot)}$ and the trace distance, we have

$$P(B_R\, \rho^\alpha_{ER}\, B_R^\dagger, B_R\, \tilde{\rho}^\alpha_{ER}\, B_R^\dagger) \leqslant d_L \sqrt{\nu} \ . \tag{C.13}$$

On the other hand, using again the joint concavity of the fidelity, we have

$$\begin{aligned} F\big(B_R\, \tilde{\rho}^\alpha_{ER}\, B_R^\dagger, \zeta_\alpha \otimes \sigma_R\big) &= F\Big(\sum\nolimits_x \rho_\alpha^{x,x} \otimes (B_R\, |x\rangle\langle x|\, B_R^\dagger), \sum\nolimits_x \zeta_\alpha \otimes (B_R\, |x\rangle\langle x|\, B_R^\dagger) \Big) \\ &\geqslant \sum_x \langle x|\, B^\dagger B\, |x\rangle_R\, F\left(\rho_\alpha^{x,x} \otimes \frac{B_R\, |x\rangle\langle x|\, B_R^\dagger}{\langle x|\, B^\dagger B\, |x\rangle_R}, \zeta_\alpha \otimes \frac{B_R\, |x\rangle\langle x|\, B_R^\dagger}{\langle x|\, B^\dagger B\, |x\rangle_R} \right) \\ &= \sum_x \langle x|\, B^\dagger B\, |x\rangle_R\, F(\rho_\alpha^{x,x}, \zeta_\alpha) \\ &\geqslant \sum_x \langle x|\, B^\dagger B\, |x\rangle_R \sqrt{1 - \epsilon'^2} \\ &\geqslant \sqrt{1 - \epsilon'^2} \ , \end{aligned} \tag{C.14}$$

recalling our assumption (C.5a) and using the fact that $\mathrm{tr}(B^\dagger B) = \mathrm{tr}(BB^\dagger) = 1$; hence

$$P\big(B_R\, \tilde{\rho}^\alpha_{ER}\, B_R^\dagger, \zeta_\alpha \otimes \sigma_R\big) \leqslant \epsilon' \ . \tag{C.15}$$

By triangle inequality for the purified distance, we have

$$\begin{aligned} P\big(B_R \rho^\alpha_{BR} B_R^\dagger, \zeta_\alpha \otimes \sigma_R\big) &\leqslant P(B_R\, \rho^\alpha_{ER}\, B_R^\dagger, B_R\, \tilde{\rho}^\alpha_{ER}\, B_R^\dagger) + P\big(B_R\, \tilde{\rho}^\alpha_{ER}\, B_R^\dagger, \zeta_\alpha \otimes \sigma_R\big) \\ &\leqslant d_L \sqrt{\nu} + \epsilon' \ . \end{aligned} \tag{C.16}$$

Returning to (C.10), we now have $F\big(B_R \rho^\alpha_{BR} B_R^\dagger, \zeta_\alpha \otimes \sigma_R\big) \geqslant \sqrt{1 - (d_L \sqrt{\nu} + \epsilon')^2}$ and hence

$$F\big(\widehat{\mathcal{N} \circ \mathcal{E}}(\sigma_{LR}), \zeta_{C'E} \otimes \sigma_R\big) \geqslant \sqrt{1 - (d_L \sqrt{\nu} + \epsilon')^2} \ . \tag{C.17}$$

As this holds for any $|\sigma\rangle_{LR}$, we deduce that

$$f(\mathcal{N} \circ \mathcal{E}) \geqslant \sqrt{1 - (d_L \sqrt{\nu} + \epsilon')^2} \ , \tag{C.18}$$

which implies (C.6). ∎

SM. D: Calculations for covariant code examples

D.1. Three-rotor secret-sharing code

Sharp cutoff. We complete the exposition in the main text in §VI A by calculating the approximation parameter $\epsilon_{\mathrm{worst}}(\mathcal{N} \circ \mathcal{E}^{(m)})$ of the constructed code.

The strategy is to apply Proposition 9. First write the operators (C.3) in our situation,

$$\rho_i^{x,x'} = \mathrm{tr}_{A\setminus A_i}(V|x\rangle\langle x'|V^\dagger) \ . \tag{D.1}$$

We need to show that $\rho_i^{x,x}$ is approximately constant of x and that $\rho_i^{x,x'}$ is very small for $x \neq x'$. The latter condition turns out to be simple: for any i and for any $x \neq x'$, we will see that $\rho_i^{x,x'} = 0$; hence we may take $\nu = 0$ in Proposition 9.

For each i, we would like to show that there exists a state ζ_i such that $\rho_i^{x,x}$ is close to ζ_i in fidelity distance for each x. We choose to work with the trace distance instead, and deduce that the states are close in fidelity using the relation $F(\cdot,\cdot) \geqslant \sqrt{1 - 2\delta(\cdot,\cdot)}$ between the fidelity and the trace distance. We bound the trace distance as follows. For each i, we find a positive semidefinite operator τ_i with the property tht $\rho_i^{x,x} \geqslant \tau_i$ for all x. This implies that $\rho_i^{x,x} = \tau_i + \Delta_i^x$ for some positive semidefinite operators Δ_i^x with $\mathrm{tr}(\Delta_i^x) = 1 - \mathrm{tr}(\tau_i)$. Define $\zeta_i = \tau_i + \xi_i$, for any freely chosen $\xi_i \geqslant 0$ with $\mathrm{tr}(\xi_i) = 1 - \mathrm{tr}(\tau_i)$. Then, we have $\rho_i^{x,x} - \zeta_i = \Delta_i^x - \xi_i$, and $\delta(\rho_i^{x,x}, \zeta_i) = (1/2)\|\rho_i^{x,x} - \zeta_i\|_1 \leqslant (1/2)(\mathrm{tr}(\Delta_i^x) + \mathrm{tr}(\xi_i)) = 1 - \mathrm{tr}(\tau_i)$. To summarize: If we find, for each i, an operator $\tau_i \geqslant 0$ with $\rho_i^{x,x} \geqslant \tau_i$ for all x, then we can deduce that there are states ζ_i such that

$$F(\rho_i^{x,x}, \zeta_i) \geqslant \sqrt{1 - \epsilon'^2} \ , \tag{D.2}$$

where $\epsilon' = \min_i \sqrt{2(1 - \mathrm{tr}(\tau_i))}$.

We may calculate the corresponding operators $\rho_i^{x,x'}$, starting with $i = 1$:

$$\begin{aligned}
\rho_1^{x,x'} &= \mathrm{tr}_{A\setminus A_1}(V|x\rangle\langle x'|V^\dagger) \\
&= \frac{1}{2m+1} \sum_{y,y'=-m}^{m} |{-3y}\rangle\langle{-3y'}| \, \delta_{y-x,y'-x'} \, \delta_{2(x+y),2(x'+y')} \\
&= \frac{\delta_{x,x'}}{2m+1} \sum_{y=-m}^{m} |{-3y}\rangle\langle{-3y}|_{A_1} \ ,
\end{aligned} \tag{D.3}$$

since the two Kronecker deltas force $x' = x$ and $y' = y$. Similarly, we have

$$\rho_2^{x,x'} = \frac{\delta_{x,x'}}{2m+1} \sum_{y=-m}^{m} |y - x\rangle\langle y - x|_{A_2} \tag{D.4}$$

$$\rho_3^{x,x'} = \frac{\delta_{x,x'}}{2m+1} \sum_{y=-m}^{m} |2(x+y)\rangle\langle 2(x+y)|_{A_3} \ . \tag{D.5}$$

First of all, for each of $i = 1, 2, 3$ we have that $\rho_i^{x,x'} = 0$ if $x \neq x'$. Then, we have that $\rho_1^{x,x}$ is already independent of x, so we may choose $\tau_1 = \rho_1^{1,1} = \rho_1^{x,x} \ \forall\, x$. Next, $\rho_2^{x,x}$ is diagonal, with constant diagonal elements $1/(2m+1)$ at states $-m-x, -m-x+1, \ldots, m-x$. We

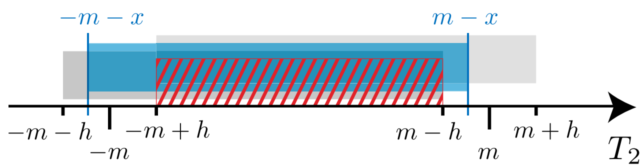

FIG. 2: Finding the "common minimal operator" for the different $\rho_2^{x,x}$'s. The solid rectangles illustrate the spectra of the different $\rho_2^{x,x}$. The eigenvalues are all equal and the rectangles are displaced vertically for readability. The hatched region corresponds to a good choice for τ_2.

may thus choose

$$\tau_2 = \frac{1}{2m+1} \sum_{u=-m+h}^{m-h} |u\rangle\langle u| \ , \tag{D.6}$$

such that $\rho_2^{x,x} \geqslant \tau_2$ for all x (Fig. 2). Finally, $\rho_3^{x,x}$ is also diagonal with elements $1/(2m+1)$ at states $-2m+2x, -2m+2x+2, \ldots, 2m+2x$. Similarly we may choose

$$\tau_3 = \frac{1}{2m+1} \sum_{u=-m+h}^{m-h} |2u\rangle\langle 2u| \ , \tag{D.7}$$

which guarantees that $\rho_3^{x,x} \geqslant \tau_3$ for each x. We have

$$\begin{aligned} \mathrm{tr}(\tau_1) &= 1 \ ; \\ \mathrm{tr}(\tau_2) &= \frac{2(m-h)+1}{2m+1} = 1 - \frac{2h}{2m+1} \ ; \\ \mathrm{tr}(\tau_3) &= 1 - \frac{2h}{2m+1} \ , \end{aligned} \tag{D.8}$$

so we may set according to the above $\epsilon' = \sqrt{4h/(2m+1)}$. According to Proposition 9, the code $V_{L\to A}^{(m)}$ is an approximate quantum error-correcting code with

$$\epsilon_{\text{worst}}(\mathcal{N} \circ \mathcal{E}^{(m)}) \leqslant \epsilon' \ . \tag{D.9}$$

We have $1/(2m+1) \approx 1/2m$, and to first order in h/m, we have

$$\epsilon(\mathcal{N} \circ \mathcal{E}^{(m)}) \lesssim \sqrt{2}\sqrt{\frac{h}{m}} \ . \tag{D.10}$$

So our codes become good in the limit $h/m \to 0$.

To compare with our bound (26), we choose $q_1 = q_2 = q_3 = 1/3$ and note that $\delta = 0$, $\eta = 0$, and $\Delta T_L = 2h$. Also, we have

$$\Delta T_1 = 2 \cdot 3m \ ; \qquad \Delta T_2 = 2(m+h) \ ; \qquad \Delta T_3 = 4(m+h) \ ; \tag{D.11}$$

so, for $m \gg h$, we have $\max_i q_i^{-1}\Delta T_i \approx 18m$. Our bound then reads

$$\epsilon(\mathcal{N}\circ\mathcal{E}) \geqslant \frac{1}{2}\frac{\Delta T_L}{\max_i q_i^{-1}\Delta T_i} \approx \frac{1}{2}\frac{2h}{18m} = \frac{1}{18}\,\frac{h}{m}\ . \tag{D.12}$$

Smooth cutoff. Again, we make use of Proposition 9. First, we compute the normalization factor as

$$c_w = \sum_{y=-\infty}^{\infty} e^{-\frac{y^2}{2w^2}} = \sum_{y=-\infty}^{\infty}\left(e^{-\frac{1}{2w^2}}\right)^{(y^2)} = \vartheta_3\Big(0, e^{-\frac{1}{2w^2}}\Big)\ , \tag{D.13}$$

where $\vartheta_3(z,q)$ is Jacobi's theta function.[2] A straightforward observation is that $c_w \geqslant 1$ (the term $y=0$ in the sum is already equal to one).

We need to determine the operators $\rho_{1,2,3}^{x,x'}$. We have

$$\begin{aligned}\rho_1^{x,x'} &= \mathrm{tr}_{A\setminus A_1}(V|x\rangle\langle x'|V^\dagger)\\ &= c_w^{-1}\sum_{y,y'=-\infty}^{\infty} e^{-\frac{y^2}{4w^2}-\frac{y'^2}{4w^2}}|{-3y}\rangle\langle{-3y'}|\,\delta_{y-x,y'-x'}\,\delta_{2(x+y),2(x'+y')}\\ &= \frac{\delta_{x,x'}}{c_w}\sum_{y=-\infty}^{\infty} e^{-\frac{y^2}{2w^2}}|{-3y}\rangle\langle{-3y}|_{A_1}\ .\end{aligned} \tag{D.14}$$

Similarly, for the second and third systems,

$$\rho_2^{x,x'} = \frac{\delta_{x,x'}}{c_w}\sum_{y=-\infty}^{\infty} e^{-\frac{(y+x)^2}{2w^2}}|y\rangle\langle y|_{A_2} \tag{D.15}$$

$$\rho_3^{x,x'} = \frac{\delta_{x,x'}}{c_w}\sum_{y=-\infty}^{\infty} e^{-\frac{(y-x)^2}{2w^2}}|2y\rangle\langle 2y|_{A_3}\ . \tag{D.16}$$

Hence, we have $\|\rho_1^{x,x'}\|_1 = \|\rho_2^{x,x'}\|_1 = \|\rho_3^{x,x'}\|_1 = 0$ for all $x \neq x'$, so the conditions (C.5b) are satisfied with $\nu = 0$.

Now we need to verify the conditions (C.5a). For the first system, $\rho_1^{x,x}$ doesn't depend on x, so choosing $\zeta_1 = \rho_1^{0,0}$ we have $P(\rho_1^{x,x},\zeta_1) = 0$ for all x. For the second system, we choose $\zeta_2 = \rho_2^{0,0}$ and calculate

$$F\big(\rho_2^{x,x},\zeta_2\big) = \sum_{y=-\infty}^{\infty}\sqrt{\frac{1}{c_w}e^{-\frac{(y+x)^2}{2w^2}}}\sqrt{\frac{1}{c_w}e^{-\frac{y^2}{2w^2}}} = \frac{1}{c_w}\sum_{y=-\infty}^{\infty} e^{-\frac{(y+x)^2+y^2}{4w^2}} \geqslant e^{-\frac{h^2}{8w^2}}\ , \tag{D.17}$$

where the calculation of the last inequality is carried out below in Lemma 10. Hence

$$P\big(\rho_2^{x,x},\zeta_2\big) \leqslant \sqrt{1-e^{-\frac{h^2}{4w^2}}} = \frac{h}{2w}\sqrt{1+O\Big(\Big(\frac{h}{w}\Big)^2\Big)} = \frac{h}{2w} + O\Big(\Big(\frac{h}{w}\Big)^3\Big)\ . \tag{D.18}$$

---

[2] See DLMF: *http://dlmf.nist.gov/20*. Our notation follows DLMF's notation.

Now, we look at the third system. Defining $\zeta_3 = \rho_3^{0,0}$, we have

$$F\big(\rho_3^{x,x}, \zeta_3\big) = \sum_{y=-\infty}^{\infty} \sqrt{\frac{1}{c_w} e^{-\frac{(y-x)^2}{2w^2}}} \sqrt{\frac{1}{c_w} e^{-\frac{y^2}{2w^2}}} = \frac{1}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y-x)^2+y^2}{4w^2}} \geqslant e^{-\frac{h^2}{8w^2}} \ , \tag{D.19}$$

invoking again the calculation in Lemma 10. Hence

$$P\big(\rho_3^{x,x}, \zeta_3\big) \leqslant \sqrt{1 - e^{-\frac{h^2}{4w^2}}} = \frac{h}{2w} + O\Big(\Big(\frac{h}{w}\Big)^3\Big) \ . \tag{D.20}$$

We are now in position to apply our criterion. Proposition 9 tells us that

$$\epsilon(\mathcal{N} \circ \mathcal{E}^{(w)}) \leqslant \sqrt{1 - e^{-\frac{h^2}{4w^2}}} = \frac{h}{2w} + O\Big(\Big(\frac{h}{w}\Big)^3\Big) \ . \tag{D.21}$$

Hence, our code's performance scales as $(1/2)(h/w)$. For instance, it performs well in the limit $h/w \to 0$, for instance in the limit $w \to \infty$ with a constant h.

Let's now see how our bound applies to our code (we need the more general bound, because we are dealing with infinite-dimensional systems with an unbounded charge observable). We need to cut off tails of the codeword states on the physical systems to make the range of charge values finite. Choose cut-offs $W_1, W_2, W_3 \geqslant 0$ for each physical system. We would like to compute an upper bound to $\sum \psi_x \psi_{x'}^* \operatorname{tr}(\Pi_i^\perp \rho_i^{x,x'}) = \sum |\psi_x|^2 \operatorname{tr}(\Pi_i^\perp \rho_i^{x,x})$, where $\Pi_i^\perp$ projects outside of the cut-off region. We have

$$\operatorname{tr}(\Pi_1^\perp \rho_1^{x,x}) = c_w^{-1} \sum_{|3y|>W_1} e^{-\frac{y^2}{2w^2}} \leqslant c_w^{-1} \sum_{|y|>\lfloor W_1/3 \rfloor} e^{-\frac{y^2}{2w^2}} \leqslant \frac{2}{c_w} \frac{w^2}{\lfloor W_1/3 \rfloor} e^{-\frac{(\lfloor W_1/3 \rfloor)^2}{2w^2}} \ , \tag{D.22}$$

where the bound is calculated in Lemma 11 below. Then,

$$\operatorname{tr}(\Pi_2^\perp \rho_2^{x,x}) = c_w^{-1} \sum_{|y|>W_2} e^{-\frac{(y+x)^2}{2w^2}} \leqslant \frac{2}{c_w} \frac{w^2}{W_2 - |x|} e^{-\frac{(W_2-|x|)^2}{2w^2}} \leqslant \frac{2}{c_w} \frac{w^2}{W_2 - h} e^{-\frac{(W_2-h)^2}{2w^2}} \ . \tag{D.23}$$

Similarly,

$$\operatorname{tr}(\Pi_3^\perp \rho_3^{x,x}) = c_w^{-1} \sum_{|2y|>W_3} e^{-\frac{(y-x)^2}{2w^2}} \leqslant \frac{2}{c_w} \frac{w^2}{\lfloor W_3/2 \rfloor - h} e^{-\frac{(\lfloor W_3/2 \rfloor - h)^2}{2w^2}} \ . \tag{D.24}$$

Hence, choosing $W_1 = W_2 = W_3 =: W$ with $W \geqslant 2h$ and choosing for simplicity W as a multiple of 6, we have $\lfloor W_1/3 \rfloor = W/3 \geqslant (1/3)(W - 2h)$, as well as $W_2 - h \geqslant W - 2h$ and also $\lfloor W_3/2 \rfloor - h = (1/2)(W - 2h)$; furthermore $W - 2h \geqslant (1/2)(W - 2h) \geqslant (1/3)(W - 2h)$. Then,

$$\begin{aligned}
\Big|\sum\nolimits_i \operatorname{tr}(\Pi_i^\perp \rho_i^{x,x})\Big| &\leqslant \frac{2}{c_w} \frac{w^2}{(1/3)(W-2h)} e^{-\frac{(\frac{1}{3}(W-2h))^2}{2w^2}} + \frac{2}{c_w} \frac{w^2}{W-2h} e^{-\frac{(W-2h)^2}{2w^2}} \\
&\quad + \frac{2}{c_w} \frac{w^2}{(1/2)(W-2h)} e^{-\frac{(\frac{1}{2}(W-2h))^2}{2w^2}} \\
&\leqslant \frac{12}{c_w} \frac{w^2}{W-2h} e^{-\frac{(W-2h)^2}{18w^2}} \leqslant \frac{12\, w^2}{W-2h} e^{-\frac{(W-2h)^2}{18w^2}} =: \eta \ ,
\end{aligned} \tag{D.25}$$

recalling that $c_w \geqslant 1$. Also, $\Delta T_i = 2W_i = 2W$ by construction. Furthermore $\Delta T_L = 2h$ and $\delta = 0$. So, our bound reads (assuming that the noise erasure probabilities are $q_1 = q_2 = q_3 = 1/3$)

$$\begin{aligned}\epsilon(\mathcal{N}\circ\mathcal{E}) &\geqslant \frac{1}{2}\,\frac{1}{(\max_i q_i^{-1})\cdot 2W}\,[2h-2\eta] = \frac{1}{2}\,\frac{1}{6W}\left[2h - \frac{24\,w^2}{W-2h}\,e^{-\frac{(W-2h)^2}{18w^2}}\right] \\ &\approx \frac{1}{2}\,\frac{1}{6W}\left[2h - \frac{24\,w^2}{W}\,e^{-\frac{W^2}{18w^2}}\right] = \frac{h}{6W} - \frac{4\,w^2}{W^2}\,e^{-\frac{W^2}{18w^2}}\ . \end{aligned} \tag{D.26}$$

considering the regime $W \gg h$, i.e., $W - 2h \approx W$. Now, if we choose the cutoff $W = \beta w$ to be proportional to w, then we can write our bound as a function of h/w:

$$\epsilon(\mathcal{N}\circ\mathcal{E}) \gtrsim \frac{1}{6\beta}\,\frac{h}{w} - \frac{4\,e^{-\beta^2/8}}{\beta^2}\ . \tag{D.27}$$

The second term is exponentially suppressed in β; so choosing β only very moderately large, we get a bound which is effectively proportional to h/w with a proportionality constant $1/(6\beta)$.

Now we find a suitable β to plug into (D.27) to get a bound in terms of h/w only. If we attempt to minimize the bound (D.27), we get as minimization condition

$$0 = \frac{\partial}{\partial\beta}\big(\text{bound}\big) = -\frac{1}{6\beta^2}\,\frac{h}{w} + \frac{8\,e^{-\beta^2/8}}{\beta^3} + \frac{e^{-\beta^2/8}}{\beta} = \frac{1}{\beta^2}\left[-\frac{h}{6w} + \frac{8+\beta^2}{\beta}\,e^{-\beta^2/8}\right]\ . \tag{D.28}$$

Writing $z = \beta^2/4$ (i.e., $\beta = 2\sqrt{z}$) we obtain $h/(6w) = e^{-z/2}\,(4+2z)/(\sqrt{z})$; the square of this equation gives

$$\frac{h^2}{36w^2} = \left[4z + 16 + \frac{16}{z}\right]e^{-z}\ . \tag{D.29}$$

To render this equation tractable, and since we only have to come up with an approximate educated guess for β, we may simplify this equation by keeping the leading term, expecting that z should be moderately large, yielding

$$\left(\frac{h}{12w}\right)^2 \approx z e^{-z}\ . \tag{D.30}$$

The solution to the equation $x^2 = ze^{-z}$ is given by the Lambert W function[3] with $z = -W(-x^2)$. Using the expansion of the negative branch W_{m} of the function near $z \to -\infty$, we have[4] $-W_{\mathrm{m}}(-x^2) \approx \ln(1/x^2)$, and hence we may select $z \approx \ln\big((12w/h)^2\big) = 2\ln(12w/h)$. This in turn yields the educated guess $\beta = 2\sqrt{2\ln(12w/h)}$ to plug into (D.27), and the bound becomes

$$\epsilon(\mathcal{N}\circ\mathcal{E}) \gtrsim \frac{h}{12w}\left[\frac{1}{\sqrt{2\ln(12w/h)}} - \frac{1}{2\ln(12w/h)}\right] \approx \frac{h/w}{12\sqrt{2\ln(w/h)}}\ , \tag{D.31}$$

using $\sqrt{\ln(12w/h)} = \sqrt{\ln(w/h) + \ln(12)} \approx \sqrt{\ln(w/h)}$.

[3] *https://dlmf.nist.gov/4.13*
[4] *https://dlmf.nist.gov/4.13.E11*

Lemma 10. *We have for integer x, h, with $|x| \leqslant h$ and with h even,*

$$\frac{1}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y\pm x)^2+y^2}{4w^2}} \geqslant e^{-\frac{h^2}{8w^2}} \ . \tag{D.32}$$

Proof of Lemma 10. First, we may assume without loss of generality that we have the "+" case in the exponent (or else simply send $x \to -x$). Completing the square, we have $(y+x)^2 + y^2 = 2y^2 + 2xy + x^2 = 2(y + x/2)^2 + x^2/2$, and hence

$$\frac{1}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y+x)^2+y^2}{4w^2}} = \frac{1}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y+x/2)^2}{2w^2} - \frac{x^2}{8w^2}} = \frac{e^{-\frac{x^2}{8w^2}}}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y+x/2)^2}{2w^2}} \ . \tag{D.33}$$

At this point we need to distinguish the case where x is even from the case where x is odd. Assuming first that x is even, we may redefine $y \to y + x/2$ in the summation and we have

$$\text{(D.33)}\ [x \text{ even}] = \frac{e^{-\frac{x^2}{8w^2}}}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{y^2}{2w^2}} = \frac{e^{-\frac{x^2}{8w^2}}}{c_w} \cdot c_w = e^{-\frac{x^2}{8w^2}} \geqslant e^{-\frac{h^2}{8w^2}} \ , \tag{D.34}$$

recalling that $|x| \leqslant h$. In the case that x is odd, we need to work a little bit more; we may redefine $y \to y + (x-1)/2$, and we have

$$\text{(D.33)}\ [x \text{ odd}] = \frac{e^{-\frac{x^2}{8w^2}}}{c_w} \sum_{y=-\infty}^{\infty} e^{-\frac{(y+\frac{1}{2})^2}{2w^2}} = \frac{e^{-\frac{x^2}{8w^2}}}{c_w}\, \vartheta_2\Big(0, e^{-\frac{1}{2w^2}}\Big) \ , \tag{D.35}$$

using another theta function corresponding to this type of summation. Lemma 12 shows that $\vartheta_2\big(0, e^{-1/(2w^2)}\big) \geqslant e^{-1/(8w^2)}\, \vartheta_3\big(0, e^{-1/(2w^2)}\big)$, and so we have

$$\text{(D.35)} \geqslant e^{-\frac{x^2+1}{8w^2}} \geqslant e^{-\frac{(|x|+1)^2}{8w^2}} \geqslant e^{-\frac{h^2}{8w^2}} \ , \tag{D.36}$$

where we have assumed that h is even, and so $|x| + 1 \leqslant h$. ■

Lemma 11. *We have, for $W \geqslant 0$,*

$$\sum_{|y|>W} e^{-\frac{(y\pm x)^2}{2w^2}} \leqslant 2\, \frac{w^2}{W - |x|}\, e^{-\frac{(W-|x|)^2}{2w^2}} \ . \tag{D.37}$$

Proof of Lemma 11. Assume $x \geqslant 0$, or else redefine $x \to -x$. We have

$$\sum_{|y|>W} e^{-\frac{(y\pm x)^2}{2w^2}} \leqslant 2 \cdot \sum_{y>W-x} e^{-\frac{y^2}{2w^2}} \leqslant 2 \cdot \sum_{y\geqslant W-x+1} e^{-\frac{y^2}{2w^2}} \leqslant 2 \int_{W-x}^{\infty} dy\, e^{-\frac{y^2}{2w^2}} \ , \tag{D.38}$$

where the integral is necessarily an overestimation of the sum, as the sum can be seen as an integral of a step function, where each step is specified at the right edge by the value of the integrand function; this step function lies beneath the actual decreasing function $e^{-y^2/(2w^2)}$.

Setting $t = y/(w\sqrt{2})$,

$$\text{(D.38)} = 2\int_{\frac{W-x}{w\sqrt{2}}}^{\infty} dt\, w\sqrt{2}\, e^{-t^2} = w\,\sqrt{2\pi}\,\frac{2}{\sqrt{\pi}}\int_{\frac{W-x}{w\sqrt{2}}}^{\infty} dt\, e^{-t^2} = w\,\sqrt{2\pi}\,\operatorname{erfc}\left(\frac{W-x}{w\sqrt{2}}\right)\ . \quad \text{(D.39)}$$

We use the known bound[5]

$$\operatorname{erfc}(z) \leqslant \frac{e^{-z^2}}{z\sqrt{\pi}}\ , \tag{D.40}$$

leading to

$$\text{(D.39)} \leqslant 2\,\frac{w^2}{W-x}\,e^{-\frac{(W-x)^2}{2w^2}}\ . \qquad \blacksquare$$

Finally, we prove a property of the theta functions that we used above.

Lemma 12. *Let* $0 < q \leqslant 1$*, and let* $z \in \mathbb{C}$ *with* $\operatorname{Re}(z) = 0$ *and* $\operatorname{Im}(z) \geqslant 0$*. Then*

$$\vartheta_3(z,q) \geqslant \vartheta_3(0,q)\ . \tag{D.41}$$

Furthermore, we have

$$\vartheta_2(0,q) \geqslant q^{1/4}\,\vartheta_3(0,q)\ . \tag{D.42}$$

Proof of Lemma 12. We start by proving (D.41). Writing $q = e^{i\pi\tau}$ with $\operatorname{Re}(\tau) = 0$ and $\operatorname{Im}(\tau) \geqslant 0$, we have[6]

$$\vartheta_3(z,q) = \vartheta_3(0,q)\cdot\prod_{n=1}^{\infty}\frac{\cos\big((n-\frac{1}{2})\pi\tau + z\big)\cos\Big((n-\frac{1}{2})\pi\tau - z\Big)}{\cos^2\big((n-\frac{1}{2})\pi\tau\big)} =: \vartheta_3(0,q)\cdot\prod_{n=1}^{\infty} a_n\ . \tag{D.43}$$

We will show that the product is greater than 1, by showing that $a_n \geqslant 1$ for each n. We have

$$a_n = \frac{\cos(i(a+b))\,\cos(i(a-b))}{\cos^2(i\,a)}\ , \tag{D.44}$$

defining $a, b \geqslant 0$ as $a = (n-1/2)\,\pi\,\operatorname{Im}(\tau)$ and $b = \operatorname{Im}(z)$. Since $\cos(i\,\varphi) = \cosh(\varphi)$, we have

$$a_n = \frac{\cosh(a+b)\,\cosh(a-b)}{\cosh^2(a)}\ . \tag{D.45}$$

(By the way, this is another way of seeing that $\vartheta_3(z,q)$ must be real and positive, since all the a_n are real positive and $\vartheta_3(0,q)$ is real positive as given by its series representation. Recall that z is pure imaginary with $\operatorname{Im}(z) \geqslant 0$, and that $0 < q \leqslant 1$.) With the usual properties of

---

[5] See for instance *http://dlmf.nist.gov/7.8.E4* or *http://mathworld.wolfram.com/Erfc.html*
[6] See *http://dlmf.nist.gov/20.5.E7*, Eq. (20.5.7)

the hyperbolic functions, we have

$$\begin{aligned}\cosh(a+b)\cosh(a-b) &= [\cosh(a)\cosh(b)+\sinh(a)\sinh(b)][\cosh(a)\cosh(b)-\sinh(a)\sinh(b)]\\ &= \cosh^2(a)\cosh^2(b)-\sinh^2(a)\sinh^2(b)\\ &= \cosh^2(a)\cosh^2(b)\big[1-\tanh^2(a)\tanh^2(b)\big]\\ &\geqslant \cosh^2(a)\cosh^2(b)\big[1-\tanh^2(b)\big]=\cosh^2(a)\ ,\end{aligned} \tag{D.46}$$

using $\tanh(a)\leqslant 1$ and $1/\cosh^2(b)=1-\tanh^2(b)$. Hence finally, $a_n\geqslant 1$. This proves (D.41).

To prove (D.42), we invoke the following property of the theta functions,[7] valid for any $q=e^{i\pi\tau}$,

$$\vartheta_2\big(0,q\big)=q^{1/4}\,\vartheta_3\left(\frac{1}{2}\pi\tau,q\right)\ . \tag{D.47}$$

For $0<q\leqslant 1$, necessarily τ is pure imaginary with $\mathrm{Im}(\tau)\geqslant 0$; we may thus invoke (D.41), which proves (D.42). ■

D.2. Five-rotor perfect code

The normalized encoding for this code is

$$|x\rangle\to\frac{1}{\sqrt{c_{w,x}}}\sum_{j,k,l,m,n\in\mathbb{Z}}e^{-\frac{1}{4w^2}\left(j^2+k^2+l^2+m^2+n^2\right)}T^{(\infty)}_{jklmnx}|j,k,l,m,n\rangle\ , \tag{D.48}$$

where $c_{w,x}$ is the normalization and $T^{(\infty)}$ is defined in Eq. (45).

Single erasure. We first calculate $\rho_\ell^{x,x}$ for $\ell\in\{1,2,3,4,5\}$ and then outline why $||\rho_\ell^{x,x'\neq x}||_1=O(e^{-cw^2})$. By the cyclic permutation symmetry of the code, we only have to calculate $\rho_1^{x,x}$. Performing the partial trace and simplifying all Kronecker delta functions leaves us with the diagonal reduced density matrix

$$\rho_1^{x,x}=\frac{1}{c_{w,x}}\sum_{j\in\mathbb{Z}}\left(\sum_{k,l,m\in\mathbb{Z}}e^{-\frac{1}{2w^2}\left(j^2+k^2+l^2+m^2+[j+k+l+m-x]^2\right)}\right)|j\rangle\langle j| \tag{D.49}$$

Now we apply the Poisson summation formula,

$$\sum_{n\in\mathbb{Z}}f(n)=\sum_{n\in\mathbb{Z}}\int_{-\infty}^{\infty}dxe^{2\pi inx}f(x)\ , \tag{D.50}$$

to each of the three sums above. Typically, the $n=0$ term on the right-hand-side is dominant (i.e., the leading order contribution in the large-w limit), and taking only this term is equivalent to approximating the sum with a Gaussian integral. Each of the remaining terms suppressed as $O(e^{-cw^2})$, where c is a positive constant increasing with n. Because c

[7] See *http://dlmf.nist.gov/20.2.E12*, Eq. (20.2.12)

increases with n, the $n+1$-th term is subleading with respect to the nth term. Thus, the entire sum of exponentially suppressed terms can itself be bounded by an exponential (e.g., $e^{-2x}+e^{-3x}<e^{-x}$ for $x>1$). We omit these corrections and focus on the dominant term $k=l=m=0$ after having applied Poisson summation to Eq. (D.49):

$$\rho_1^{x,x} \sim \frac{\sqrt{2}\pi^{3/2}w^3}{c_{w,x}} \sum_{j\in\mathbb{Z}} e^{-\frac{5j^2-2jx+x^2}{8w^2}} |j\rangle\langle j| \,. \tag{D.51}$$

Forcing $\mathrm{tr}\{\rho_1^{x,x}\}=1$ and once again approximating the resulting sum with an integral solves for the normalization $c_{w,x}$ in the large w limit. Plugging that back into the above equation and simplifying produces

$$\rho_1^{x,x} \sim \sqrt{\frac{5}{2\pi}}\frac{1}{2w} \sum_{j\in\mathbb{Z}} e^{-\frac{(x-5j)^2}{40w^2}} |j\rangle\langle j| \,. \tag{D.52}$$

Now we calculate the fidelity of the above state to $\rho_1^{0,0}$. Using the fact that the states commute with each other, taking the square root of each entry in the resulting diagonal matrix, and applying Poisson summation yields

$$F^2\left(\rho_1^{x,x},\rho_1^{0,0}\right) = \sqrt{\frac{5}{2\pi}}\frac{1}{2w} \sum_{j\in\mathbb{Z}} e^{-\frac{50j^2-10jx+x^2}{80w^2}} \sim e^{-\frac{x^2}{160w^2}} \geqslant e^{-\frac{h^2}{160w^2}} \,. \tag{D.53}$$

Plugging this into the infidelity yields the result (46).

Returning to the $x\neq x'$ case, we show why those cases do not significantly contribute. The reduced density matrix is of the form

$$\rho_1^{x,x'} = \frac{1}{\sqrt{c_{w,x}c_{w,x'}}} \sum_{j\in\mathbb{Z}} \left(\sum_{k,l,m\in\mathbb{Z}} \gamma_{j,k,l,m}^{x,x'} \right) |j\rangle\langle j+x-x'|$$

where $\gamma_{j,k,l,m}^{x,x'}$ is a product of a Gaussian in the variables k,l,m (just like the $x=x'$ case above) and a phase $\propto 2\pi\Phi$ (which goes away when $x=x'$). We first apply Poisson summation to the internal three sums and evaluate the normalizations in the large-w limit. In this case, the centers of the Gaussians in the k,l,m-sum depend on Φ and the dominant term on the right-hand-side of Eq. (D.50) may not longer be the center-of-mass term $n=0$. We will however set Φ to be an irrational number close to zero from now on, i.e., taking $\Phi\ll 1$ while making sure that $\Phi w\to\infty$. This makes sure that the center-of-mass mode is dominant. Writing the norm and applying Poisson summation to the remaining sum reveals

$$\begin{aligned} \|\rho_1^{x,x'}\|_1 &\sim \sqrt{\frac{5}{2\pi}}\frac{1}{2w} e^{-2\pi^2\Phi^2w^2\left(x-x'\right)^2} \sum_{j\in\mathbb{Z}} e^{-\frac{25j^2-30jx+20jx'+13x^2-20xx'+8x'^2}{40w^2}} \\ &= O\left(e^{-2\pi^2\left(x-x'\right)^2\Phi^2w^2}\right) \,. \end{aligned} \tag{D.54}$$

We see that the one-norm is exponentially suppressed in w^2 for the off-diagonal (i.e., $x\neq x'$) reduced matrices.

Two erasures. We first calculate $\rho^{x,x}_{\ell,\ell'}$ for $\ell,\ell' \in \{1,2,3,4,5\}$ and then argue that $||\rho^{x,x'\neq x}_{\ell,\ell'}||_1 = O(e^{-cw^2})$. Due to the cyclic permutation symmetry, we only need to calculate $\rho^{x,x}_{1,2}$ and $\rho^{x,x}_{1,3}$. Performing the partial trace, simplifying the Kronecker delta functions, plugging in the normalization, and applying Poisson summation yields

$$\rho^{x,x}_{1,2} \sim \sqrt{\frac{5}{3}}\frac{1}{2\pi w^2}\sum_{j,k\in\mathbb{Z}} e^{-\frac{10j^2+5jk+10k^2-5(j+k)x+x^2}{15w^2}}|j,k\rangle\langle j,k| \sim \rho^{x,x}_{1,3}\,. \tag{D.55}$$

In other words, both $\rho^{x,x}_{1,2}$ and $\rho^{x,x}_{1,3}$ are identical in the large w limit. Note that, unlike $\rho^{x,x}_1$, these matrices have off-diagonal elements that are exponentially suppressed in w^2. These elements have been ignored above, but we mention them in the $x \neq x'$ case below. Taking the fidelity between $\rho^{x,x}_{\ell,\ell'}$ and $\rho^{0,0}_{\ell,\ell'}$ as before yields the result $F^2(\rho^{x,x}_{1,2}, \rho^{0,0}_{1,2}) \geqslant e^{-\frac{h^2}{60w^2}}$ claimed in (47).

The $x \neq x'$ case is more difficult this time because the unapproximated reduced density matrix no longer has just one nonzero diagonal. Without any approximations, it is

$$\rho^{x,x'}_{1,2} = \sqrt{\frac{5}{3}}\frac{1}{2\pi w^2}\sum_{j,j',k\in\mathbb{Z}}\left(\sum_{l,m\in\mathbb{Z}}\gamma^{x,x'}_{j,j',k,l,m}\right)|j\rangle\langle j'|\otimes|k\rangle\langle k+x'-x+j-j'|\,. \tag{D.56}$$

Applying Poisson summation to the internal two sums for $x \neq x'$ reveals that all matrix elements are exponentially suppressed with w^2,

$$\sum_{l,m\in\mathbb{Z}}\gamma^{x,x'}_{j,j',k,l,m} = \sqrt{\frac{5}{3}}\frac{1}{2\pi w^2}O\left(e^{-\frac{4}{3}\pi^2\left[3\left(j'-j\right)^2-3\left(j'-j\right)\left(x'-x\right)+\left(x'-x\right)^2\right]\Phi^2w^2}\right)\,. \tag{D.57}$$

However, there are particular values of (j,j') for which the function in the exponent above is minimized; we select those and show that the trace norm is exponentially suppressed in w^2. For even $x-x'$, the band at $j' = j - \frac{x-x'}{2}$ decays the slowest. Ignoring all other bands and calculating the trace norm yields

$$\|\rho^{x,x'}_{1,2}\|_1 = O\left(e^{-\frac{1}{3}\pi^2\left(x-x'\right)^2\Phi^2w^2}\right)\,. \tag{D.58}$$

For odd $x-x'$, there are two bands $j' = j - \frac{x-x'\pm1}{2}$ whose entries decay the slowest. Calculating the square root of $\rho^{x,x'}_{1,2}\rho^{x,x'\dagger}_{1,2}$ is more difficult since the resulting matrix is tri-diagonal. However, ignoring the off-diagonal entries, taking the square root, and bounding the resulting integral still yields exponential scaling with w^2.

D.3. Thermodynamic codes

Here, we carry out the calculations that are relevant for §VI C of the main text.

The operators $\rho^{m,m}_d$ (reduced states on d consecutive sites) are provided as:

$$\rho^{m,m}_d = \sum_{r=-d}^{d} K^N_{r,d,m}\,|h^d_r\rangle\langle h^d_r|_d\,, \tag{D.59}$$

with

$$K^N_{r,d,m} = \frac{\binom{d}{d/2+r/2}\binom{N-d}{(N-d)/2+(m-r)/2}}{\binom{N}{N/2+m/2}} \ . \tag{D.60}$$

The fidelity between two states which commute reduces to the Bhattacharyya coefficient (the classical version of the fidelity):

$$F(\rho_d^{m,m}, \rho_d^{0,0}) = \sum_{r=-d}^{d} \sqrt{K^N_{r,d,m}}\sqrt{K^N_{r,d,0}} \ . \tag{D.61}$$

The complicated calculation is deferred to Lemma 13 below, which gives us:

$$F(\rho_d^{m,m}, \rho_d^{0,0}) \geqslant 1 - O(N^{-2}) \ . \tag{D.62}$$

This, in turn, tells us that

$$P(\rho_d^{m,m}, \rho_d^{0,0}) \leqslant O(N^{-1}) \ . \tag{D.63}$$

The "logical off-diagonal" terms $\rho_d^{m,m'}$ for $m \neq m'$ are exactly zero, because we made sure to space out the codewords in magnetization by $2d+1$, following the construction of ref. [8].

Hence, applying Proposition 9, we see that our code is an AQECC against the erasure of d consecutive sites, with

$$\epsilon(\mathcal{N} \circ \mathcal{E}) \leqslant O(N^{-1}) \ . \tag{D.64}$$

This matches exactly the scaling of our bound (26).

Lemma 13. *There exists a constant $D_{d,m}$ of N such that (for constant d, m):*

$$F(\rho_d^{m,m}, \rho_d^{0,0}) \geqslant 1 - \frac{D_{d,m}}{N^2} + O(N^{-3}) \ . \tag{D.65}$$

Proof of Lemma 13. We use Stirling's formula up to order $1/N^2$:

$$\ln(N!) = N\ln(N) - N + \frac{1}{2}\ln(2\pi N) + \frac{1}{12N} + O(N^{-3}) \ , \tag{D.66}$$

(noting that there is in fact no term of order $1/N^2$). Now, for any x, ignoring terms of order $O(N^{-3})$, we have:

$$\begin{aligned}
\ln\binom{N}{N/2+x/2} &= N\ln(N) + \frac{\ln(2\pi)}{2} + \frac{\ln(N)}{2} + \frac{1}{12N} + O(N^{-3}) \\
&\quad - \left[\left(\frac{N}{2}+\frac{x}{2}\right)\ln\left(\frac{N}{2}+\frac{x}{2}\right) + \frac{\ln(2\pi)}{2} + \frac{1}{2}\ln\left(\frac{N}{2}+\frac{x}{2}\right) + \frac{1}{6(N+x)}\right] \\
&\quad - \left[\left(\frac{N}{2}-\frac{x}{2}\right)\ln\left(\frac{N}{2}-\frac{x}{2}\right) + \frac{\ln(2\pi)}{2} + \frac{1}{2}\ln\left(\frac{N}{2}-\frac{x}{2}\right) + \frac{1}{6(N-x)}\right] \ .
\end{aligned}$$

Using the expansions

$$\ln\Big(\frac{N}{2}\pm\frac{x}{2}\Big)=\ln(N)-\ln(2)+\ln\Big(1\pm\frac{x}{N}\Big)$$
$$=\ln(N)-\ln(2)\pm\frac{x}{N}-\frac{x^2}{2N^2}\pm\frac{x^3}{3N^3}+O\big(N^{-4}\big)\ ; \tag{D.67}$$
$$\frac{1}{N\pm x}=\frac{1}{N}\left(\frac{1}{1\pm x/N}\right)=\frac{1}{N}\mp\frac{x}{N^2}+O(N^{-3})\ , \tag{D.68}$$

one continues, still keeping all the terms up to order $1/N^2$:

$$\ln\binom{N}{N/2+x/2}=N\ln(2)-\frac{\ln(N)}{2}+\ln\Big(\frac{2}{\sqrt{2\pi}}\Big)-\frac{1}{4N}-\frac{x^2}{2N}+\frac{x^2}{2N^2}+O(N^{-3})\ .$$

Now we may apply this to calculate $\ln\Big(\sqrt{K^N_{r,d,m}K^N_{r,d,0}}\Big)$, using the fact that $\ln(N-d)=\ln(N)+\ln(1-d/N)=\ln(N)-d/N-d^2/(2N^2)+O(N^{-3})$:

$$\begin{aligned}
&\ln\sqrt{K^N_{r,d,m}K^N_{r,d,0}}\\
&=\ln\binom{d}{d/2+r/2}+\frac{1}{2}\ln\binom{N-d}{(N-d)/2+(m-r)/2}\\
&\quad+\frac{1}{2}\ln\binom{N-d}{(N-d)/2-r/2}-\frac{1}{2}\ln\binom{N}{N/2+m/2}-\frac{1}{2}\ln\binom{N}{N/2}\\
&=\ln\binom{d}{d/2+r/2}-d\ln(2)\\
&\quad+\frac{1}{2}\Big\{-\frac{1}{2}\Big[\ln(N)-\frac{d}{N}-\frac{d^2}{2N^2}\Big]-\frac{1}{4}\Big(\frac{1}{N}+\frac{d}{N^2}\Big)-\frac{(m-r)^2}{2}\Big(\frac{1}{N}+\frac{d}{N^2}\Big)+\frac{(m-r)^2}{2N^2}\Big\}\\
&\quad+\frac{1}{2}\Big\{-\frac{1}{2}\Big[\ln(N)-\frac{d}{N}-\frac{d^2}{2N^2}\Big]-\frac{1}{4}\Big(\frac{1}{N}+\frac{d}{N^2}\Big)-\frac{r^2}{2}\Big(\frac{1}{N}+\frac{d}{N^2}\Big)+\frac{r^2}{2N^2}\Big\}\\
&\quad-\frac{1}{2}\Big\{-\frac{1}{2}\ln(N)-\frac{1}{4N}-\frac{m^2}{2N}+\frac{m^2}{2N^2}\Big\}-\frac{1}{2}\Big\{-\frac{1}{2}\ln(N)-\frac{1}{4N}\Big\}+O(N^{-3})\\
&=\ln\Big(2^{-d}\binom{d}{d/2+r/2}\Big)+\frac{d}{2N}+\frac{A_{m,r}}{N}+\frac{B_{d,m,r}}{N^2}+O(N^{-3})\ ,
\end{aligned}$$

with

$$A_{m,r}=\frac{1}{2}\,r(m-r)\ ; \tag{D.69a}$$
$$B_{d,m,r}=\frac{1}{4}\left[d^2-d\big(1+m^2+2r^2-2mr\big)+2r^2-2mr\right]\ . \tag{D.69b}$$

Using $0\leqslant|r|\leqslant d$, write

$$\begin{aligned}
B_{d,m,r}&\geq\frac{1}{4}\big[d^2-d\big(1+m^2+2d^2+2|m|d\big)-2|m|d\big]\\
&\geq\frac{1}{4}\big[-2d^3-d^2(2|m|-1)-d\big(1+m^2+2|m|\big)\big]\ \ =:\ \ -\frac{1}{4}\,C_{d,m}\ .
\end{aligned}$$

Then,

$$
\begin{aligned}
F(\rho_d^{m,m}, \rho_d^{0,0}) &\geqslant e^{d/(2N)}\, 2^{-d} \sum_{r=-d}^{d} \binom{d}{d/2 + r/2} \exp\left\{ \frac{A_{m,r}}{N} + \frac{B_{d,m,r}}{N^2} + O(N^{-3}) \right\} \\
&\geqslant \exp\left\{ -\frac{C_{d,m}}{4N^2} + O(N^{-3}) \right\} e^{d/(2N)}\, 2^{-d} \sum_{r=-d}^{d} \binom{d}{d/2 + r/2} \exp\left\{ \frac{A_{m,r}}{N} \right\} .
\end{aligned}
\tag{D.70}
$$

Recall the identities

$$
\sum_{r=-d}^{d} \binom{d}{d/2 + r/2} = \sum_{k=0}^{d} \binom{d}{k} = 2^d \; ; \tag{D.71a}
$$

$$
\sum_{k=0}^{d} \binom{d}{k} k = d\, 2^{d-1} \; ; \tag{D.71b}
$$

$$
\sum_{k=0}^{d} \binom{d}{k} k^2 = (d + d^2)\, 2^{d-2} \; . \tag{D.71c}
$$

We have $\exp\{A_{m,r}/N\} = 1 + r(m-r)/(2N) + r^2(m-r)^2/(8N^2) + O(N^{-3}) \geqslant 1 + r(m-r)/(2N) + O(N^{-3})$. Replacing the summation index r by $k = (d+r)/2 = 0, 1, \ldots, d$, we calculate

$$
\begin{aligned}
&\frac{2^{-d}}{2N} \sum_{r=-d}^{d} \binom{d}{d/2 + r/2} r(m-r) \\
&= \frac{2^{-d}}{2N} \sum_{k=0}^{d} \binom{d}{k} (2k - d)(m - 2k + d) \\
&= \frac{2^{-d}}{2N} \left[(-d\, m - d^2) \sum_{k=0}^{d} \binom{d}{k} + (2m + 2d + 2d) \sum_{k=0}^{d} \binom{d}{k} k - 4 \sum_{k=0}^{d} \binom{d}{k} k^2 \right] \\
&= \frac{2^{-d}}{2N} \left[(-d\, m - d^2)\, 2^d + (2m + 2d + 2d)\, \frac{d}{2}\, 2^d - 4 \frac{d + d^2}{4}\, 2^d \right] \\
&= -\frac{d}{2N} \; ,
\end{aligned}
$$

and then

$$
\begin{aligned}
2^{-d} \sum_{r=-d}^{d} \binom{d}{d/2 + r/2} \exp\left\{ \frac{A_{m,r}}{N} \right\} &\geq 2^{-d} \sum_{r=-d}^{d} \binom{d}{d/2 + r/2} \left(1 + \frac{r(m-r)}{2N} + O(N^{-3}) \right) \\
&= 1 - \frac{d}{2N} + O(N^{-3}) \; .
\end{aligned}
$$

Finally, plugging into (D.70) gives us

$$
F(\rho_d^{m,m}, \rho_d^{0,0}) \geqslant \left\{ 1 - \frac{C_{d,m}}{4N^2} + O(N^{-3}) \right\} \left\{ 1 + \frac{d}{2N} + \frac{d^2}{8N^2} + O(N^{-3}) \right\} \left\{ 1 - \frac{d}{2N} + O(N^{-3}) \right\}
$$

$$\geqslant 1 - \frac{C_{d,m}}{4N^2} - \frac{d^2}{8N^2} + O(N^{-3}) \ , \tag{D.72}$$

so we may define $D_{d,m} = C_{d,m}/4 + d^2/8$, proving the claim. ■

SM. E: Proof of the approximate Eastin-Knill theorem for quantum computation

E.1. Equivalence of the existence of a universal transversal gate set and the $U(d_L)$-covariance property of the code

First, we show that the setting of the Eastin-Knill theorem is equivalent to studying the $U(d_L)$-covariance property of the corresponding code. More precisely, we show that given a code $V_{L\to A}$, if there exists a mapping u of logical unitaries U_L to transversal physical unitaries $u(U_L) = U_1 \otimes \cdots \otimes U_n$ satisfying $V^\dagger u(U_L) V = U_L$ for all U_L, where u does not even have to be continuous, then the code $V_{L\to A}$ is necessarily covariant with respect to the full unitary group on the logical space. This holds even if the mapping $u(U_L)$ is approximate, e.g., if the physical unitary is compiled using a set of generating gates, if such mappings exist for arbitrary good precision.

The statement is pretty intuitive, because given any rule that maps logical unitaries to physical transversal unitaries, we can compose the physical unitaries corresponding to different logical unitaries, and presumably generate a *bona fide* representation by starting from a minimal generating set of unitaries. This intuition proves correct, though it is not immediately clear if the mapping generated in this way is continuous. Here we provide a derivation that smooths out these technical details.

Proposition 14. *Let $V_{L\to A}$ be any code, with $A = A_1 \otimes \cdots \otimes A_n$. Suppose that for each unitary U_L on L there exists a transversal unitary $U_A = u(U_L) = u_1(U_L) \otimes \cdots \otimes u_n(U_L)$ such that $V^\dagger u(U_L) V = U_L$ for all U_L. Then there exists a mapping u' that maps any U_L to a transversal physical unitary $u'(U_L) = u'_1(U_L) \otimes \cdots \otimes u'_n(U_L)$ such that*

(i) u' is continuous;

(ii) for all U_L, $V^\dagger u'(U_L) V = V^\dagger u(U_L) V = U_L$; and

(iii) for any U_L, U'_L, we have $u'(U_L U'_L) = u'(U_L)\, u'(U'_L)$.

In other words, there exists a tensor product representation of $U(d_L)$ on A with respect to which V is covariant.

Corollary 15. *Let $V_{L\to A}$ be any code, with $A = A_1 \otimes \cdots \otimes A_n$. Let $\{G_x\}$ be a set of generating gates of $SU(d_L)$ (which can be discrete or continuous), and suppose that for each G_x there exists a transversal implementation $U_A(G_x) = U_1(G_x) \otimes \cdots \otimes U_n(G_x)$ of the gate: $V^\dagger U_A(G_x) V = G_x$. Then there exists a tensor product representation of $SU(d_L)$ on $A_1 \otimes \cdots \otimes A_n$ with respect to which $V_{L\to A}$ is covariant.*

Proof of Corollary 15. Because $\{G_x\}$ is a generating gate set we have that for any $\epsilon > 0$ and for any unitary U_L there exists a transversal unitary $U_A^{(\epsilon)}$ such that $\|V^\dagger U_A^{(\epsilon)} V - U_L\|_\infty \leqslant \epsilon$. Compactness of the set of transversal unitaries implies that $\lim_{\epsilon\to 0^+} U_A^{(\epsilon)}$ is again transversal, and Proposition 14 applies. The difference between $U(d_L)$ and $SU(d_L)$ is a global phase which is generated by the identity operator. ■

Proof of Proposition 14. Observe first of all that any U_L, $u(U_L)$ fixes the code space $\Pi = VV^\dagger$ because $u(U_L)$ implements a logical unitary. Hence, we must necessarily have $[u(U_L), \Pi] = 0$ for all U_L. Now let $\mathcal{G}$ be the set of all transversal logical unitaries:

$$\mathcal{G} = \{U_A = U_1 \otimes \cdots \otimes U_n \ : \ [U_A, \Pi] = 0\} \ . \tag{E.1}$$

Eastin and Knill show that $\mathcal{G}$ is a Lie group [17]. This means that any $U_A \in \mathcal{G}$ can be written as $U_A = e^{iT_A} = e^{i(T_1+\cdots+T_n)}$, where T_i is a Hermitian operator acting on the i-th physical subsystem, and where $[T_A, \Pi] = 0$. Consider the set $\mathcal{G}' = V^\dagger \mathcal{G} V$, which is the projection of $\mathcal{G}$ onto the logical space. It is isomorphic (through V) to a closed subgroup of $\mathcal{G}$, so $V^\dagger \mathcal{G} V$ is again a Lie group.

By assumption, $\mathcal{G}'$ must coincide with the full unitary group $U(d_L)$ on the logical space. Indeed, all the elements of the former are unitaries on the logical space and therefore included in $U(d_L)$, and conversely, for any logical unitary $U_L \in U(d_L)$ there exists by assumption an element of $U_A \in \mathcal{G}$ such that $V^\dagger U_A V = U_L$.

The Lie algebra $\mathfrak{g}'$ of $\mathcal{G}'$ is the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$ projected on the logical space:

$$\mathfrak{g}' = \{V^\dagger T_A V \ : \ T_A \in \mathfrak{g}\} \ ; \qquad \mathfrak{g} = \{T_A = T_1 + \cdots + T_n \ : \ [T_A, \Pi] = 0\} \ , \tag{E.2}$$

where T_i is a Hermitian operator acting on the i-th physical subsystem. Now because $\mathcal{G}' = U(d_L)$, we also have $\mathfrak{g}' = \mathfrak{u}(d_L)$, recalling that the Lie algebra $\mathfrak{u}(d_L)$ of $U(d_L)$ consists of all Hermitian matrices of dimension d_L. Now let $T_L \in \mathfrak{u}(d_L) = \mathfrak{g}'$. Then there exists $T_A = T_1 + \cdots + T_n \in \mathfrak{g}$ such that $T_L = V^\dagger T_A V$. This means that for any generator T_L of the logical unitaries, there exists a corresponding generator $T_A(T_L)$ on the physical systems that commutes with the code space and that is a sum of local terms. Now choose a basis $\{T_L^{(i)}\}$ of the Lie algebra of $U(d_L)$. We can then define the mapping $u'(U_L)$ as the Lie group representation of $U(d_L)$ on the physical system A generated by the operators $T_A(T_L^{(i)})$ that span its corresponding Lie algebra. The mapping u' satisfies all the stated requirements. ∎

E.2. Proof of the approximate Eastin-Knill bound

Recall that each irrep of $U(d_L)$ is represented by a *Young diagram* λ, where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{d_L})$, and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{d_L} = 0$, and $\lambda_i \in \mathbb{Z}$. The dimension of each irrep is given by the Weyl dimension formula, which for the $U(d_L)$ group is equal to the Schur polynomial S_λ evaluated at the vector $(1, 1, \cdots, 1)$. More explicitly, it can be evaluated to

$$D_\lambda = \prod_{1 \leqslant i < j \leqslant d_L} \frac{\lambda_i - \lambda_j + j - i}{j - i} \ . \tag{E.3}$$

To derive Theorem VI we need to first prove few intermediate results. First, we will prove a bound on D_λ, based on λ_1.

Lemma 16. *The symmetric representation has the minimum dimension among the repre-*

sentations with fixed λ_1. More precisely, the following inequality holds,

$$D_{\mathrm{Sym}^{\lambda_1}} = \binom{d_L - 1 + \lambda_1}{d_L - 1} \leqslant D_\lambda \ , \tag{E.4}$$

where λ is a representation of $U(d_L)$ and $D_{\mathrm{Sym}^{\lambda_1}}$ is the dimension of the symmetric representation with the Young diagram $\lambda = (\lambda_1, 0, 0, \cdots, 0)$.

Proof of Lemma 16. Suppose that $\lambda_1 = l$. We use the dimension formula Eq. (E.3). Consider the logarithm of the dimension, which is (up to a fixed constant) equal to:

$$f(\lambda_2, \cdots, \lambda_{d_L-1}) = \sum_{1 \leqslant i < j \leqslant d_L} \log(\lambda_i - \lambda_j + j - i) \ . \tag{E.5}$$

Note that we fix $\lambda_1 = l$ and $\lambda_{d_L} = 0$, so they do not appear as parameters of f. Also, the vector $\hat{\lambda} = (\lambda_2, \cdots, \lambda_{d_L-1})$ is an integer vector in the simplex Δ with d_L extremal points $\hat{v}_i \in \mathbb{R}^{d_L-1}$, where $\hat{v}_0 = (0, 0, \cdots, 0)$, $\hat{v}_1 = (l, 0, \cdots, 0)$,$\cdots$, and $\hat{v}_{d_L-1} = (l, l, \cdots, l)$.

We first extend the function f to all of the real points in Δ, and show that f is a concave function inside Δ. This would show that the minimum of f is attained at one of its extremal points.

A direct computation of the Hessian of f, reveals that for $2 \leqslant r, s \leqslant d_L - 1$,

$$H_{r,s} = \delta_{rs} \left[- \sum_{1 \leqslant i \leqslant d_L, i \neq s} K_{is} \right] + (1 - \delta_{rs}) K_{rs} \ , \tag{E.6}$$

where $K_{rs} = 1/(\lambda_s - \lambda_r + r - s)^2$. One can see that if $w = \sum_{2 \leqslant i \leqslant d_L - 1} \alpha_i e_i$ is an arbitrary vector, then

$$w^\dagger H w = - \left(\sum_{2 \leqslant i \leqslant d_L - 1} |\alpha_i|^2 (K_{1i} + K_{id_L}) + \sum_{2 \leqslant i < j \leqslant d_L - 1} K_{ij} |\alpha_i - \alpha_j|^2 \right) \ . \tag{E.7}$$

This is a negative number, and shows that f is strictly concave. Therefore, the minimum of f is attained on one of the extremal point $\hat{v}_i$, $0 \leqslant i \leqslant d_L - 1$. Using the Weyl-dimension formula we have,

$$f(\hat{v}_i) = \prod_{j=0}^{i} \frac{\binom{l + d_L - 1 - j}{l}}{\binom{l+j}{l}} \ . \tag{E.8}$$

One can easily see that $f(\hat{v}_i)$ is increasing for $i \leqslant (d_L - 1)/2$ and decreasing for $i \geqslant (d_L - 1)/2$. Therefore, its minimum is attained at $f(\hat{v}_0) = f(\hat{v}_{d_L-1})$. ∎

Consider a fixed element in the Cartan subalgebra of $\mathfrak{su}(d_L)$, a $d_L \times d_L$ matrix $T = \mathrm{diag}(1, 0, 0, 0, \cdots, -1)$, and T_λ, the corresponding generator in the representation given by the Young diagram λ. We have the following lemma:

Lemma 17. *It holds that $\|T_\lambda\|_\infty \leqslant \lambda_1$.*

Proof of Lemma 17. A basis for the representation λ is given by different semi-standard fillings of the Young diagram λ with numbers $1 \cdots d_L$. If we indicate fillings of the λ by m_λ, then $\{|m_\lambda\rangle\}$ forms a basis for the representation λ. Although this is not an orthogonal basis, if the number content of m_λ and m'_λ are different then $|m_\lambda\rangle$ and $|m'_\lambda\rangle$ are orthogonal. This basis diagonalizes T_λ.

In particular, if $\#_i m_\lambda$ indicates the number of times that i appears in the filling m_λ, then $\langle m_\lambda|T_\lambda|m_\lambda\rangle = \#_1 m_\lambda - \#_d m_\lambda$. This immediately leads to the conclusion that the eigenvalues of T_λ are $\#_1 m_\lambda - \#_d m_\lambda$, for different fillings m_λ.

For any semi-standard filling of the Young diagrams, the numbers are strictly increasing in the columns. Therefore, $\#_i m_\lambda \leqslant \lambda_1$, as there are no repeats in the columns. So we showed that eigenvalues of T_λ are between $-\lambda_1$ and λ_1, which completes the proof. ■

We are now in a position to prove Theorem VI. We first prove the inequality with the binomial coefficent as Theorem 18, and we then specialize this bound to the different related bounds presented in §VII A.

Theorem 18 (First inequality of Theorem VI in the main text)**.** *Let $V_{L\to A}$ be an isometry that is covariant with respect to the full $SU(d_L)$ group on the logical space, and write $\mathcal{E}(\cdot) = V(\cdot)V^\dagger$. Consider the single erasure noise model represented by $\mathcal{N}$ in (10) with equal erasure probabilities, $q_i = 1/n$ for all i. Then*

$$\max_i d_i \geqslant \binom{d_L - 1 + \left\lceil [2n\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E})]^{-1}\right\rceil}{d_L - 1} . \tag{E.9}$$

In terms of the average entanglement fidelity measure, the bound reads instead

$$\max_i d_i \geqslant \binom{d_L - 1 + \left\lceil [n d_L \epsilon_{\mathrm{e}}(\mathcal{N}\circ\mathcal{E})]^{-1}\right\rceil}{d_L - 1} . \tag{E.10}$$

The bound (E.9) enables us to derive a simple bound in the event that $d_L = d_i$, as in the examples given in the main text. The binomial coefficient $\binom{a+b}{b}$ is increasing in b, which can be seen using the recurrence relation $\binom{a+b+1}{b+1} = \frac{a+b+1}{b+1}\binom{a+b}{b} \geqslant \binom{a+b}{b}$. Also, the binomial coefficient $\binom{a+b}{b}$ for $b \geqslant 2$ satisfies $\binom{a+b}{b} \geqslant \binom{a+2}{2} = (a+2)(a+1)/2 \geqslant a+2$ (assuming $a \geqslant 1$). Hence, if $d_L = d_i$, then condition (E.9) implies that $\lceil (2n\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}))^{-1}\rceil \leqslant 1$, because otherwise we would have $\binom{d_L - 1 + \lceil (2n\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}))^{-1}\rceil}{d_L - 1} \geqslant \binom{d_L - 1 + 2}{2} \geqslant d_L + 1$. This implies that, for $d_L = d_i$, we must have $\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant 1/(2n)$.

Proof of Theorem 18. Combining Lemma 16 and Lemma 17, we get

$$D_\lambda \geqslant \binom{d_L - 1 + \lceil \|T_\lambda\|_\infty \rceil}{d_L - 1} \tag{E.11}$$

Now, we return to the original problem of approximate Eastin-Knill theorem, where the group $SU(d_L)$ acts on physical subsystems. We fix the generator $T = \mathrm{diag}(1, 0, 0, 0, \cdots, -1)$ of $\mathfrak{su}(d_L)$, and let T_i be the corresponding generator acting on the subsystem i. Let $T_i = \bigoplus_\lambda T_\lambda$ be the decomposition of T_i with respect to the decomposition of the representation on subsystem i, and assume that $\hat\lambda(i)$ is the Young diagram in this direct sum with the largest $\|T_\lambda\|_\infty$. Therefore, $\|T_i\|_\infty = \left\|T_{\hat\lambda(i)}\right\|_\infty$, and we have:

$$d_i \geqslant D_{\hat\lambda(i)} \geqslant \binom{d_L - 1 + \lceil \|T_{\hat\lambda(i)}\|_\infty \rceil}{d_L - 1} = \binom{d_L - 1 + \lceil \|T_i\|_\infty \rceil}{d_L - 1} . \tag{E.12}$$

This implies

$$\max_i d_i \geqslant \max_i \binom{d_L - 1 + \lceil \|T_i\|_\infty \rceil}{d_L - 1} = \binom{d_L - 1 + \lceil \max_i \|T_i\|_\infty \rceil}{d_L - 1} . \tag{E.13}$$

Let i' denote the index of the subsystem that maximizes $\|T_i\|_\infty$, such that our bound (26) in the main text with $\Delta T_L = 2$ and $\Delta T_i \leqslant 2\|T_i\|_\infty$ reads

$$\epsilon_{\text{worst}} \geqslant \frac{1}{2n\|T_{i'}\|_\infty} , \tag{E.14}$$

noting that $\max_i \Delta T_i \leqslant 2 \max_i \|T_i\|_\infty = 2\|T_{i'}\|_\infty$, and writing $\epsilon_{\text{worst}} = \epsilon_{\text{worst}}(\mathcal{N} \circ \mathcal{E})$ as a shorthand. Therefore, $\|T_{i'}\|_\infty \geqslant (2n\epsilon_{\text{worst}})^{-1}$, and we obtain

$$\max_i d_i \geqslant \binom{d_L - 1 + \lceil (2n\epsilon_{\text{worst}})^{-1} \rceil}{d_L - 1} . \tag{E.15}$$

If we had used the bound (27) instead of (26) along with $s(T_L) = s'(T_L) = \|T_L\|_1/d_L = 2/d_L$, we would have instead of (E.14) that

$$\epsilon_{\text{e}} \geqslant \frac{1}{nd_L\|T_{i'}\|_\infty} , \tag{E.16}$$

and thus $\|T_i'\|_\infty \geqslant (\epsilon_{\text{e}} n d_L)^{-1}$. Inserting into (E.13) proves (E.10). ■

Corollary 19 (Second inequality of Theorem VI in the main text)**.** *Consider the setting in Theorem 18. Then*

$$\max_i \ln d_i \geqslant \frac{\ln(d_L - 1)}{2n\epsilon_{\text{worst}}} - \frac{\ln\big(1 + (2n\epsilon_{\text{worst}})^{-1}\big)}{2n\epsilon_{\text{worst}}} . \tag{E.17}$$

Corollary 20 (Equations (55) and (56) in the main text)**.** *Consider the setting in Theorem 18. Then*

$$\max_i \ln d_i \geqslant \begin{cases} (d_L - 1)\ln\left(\dfrac{1}{2\epsilon_{\text{worst}} n d_L}\right) & \text{(E.18a)} \\[2ex] (d_L - 1)\ln\left(\dfrac{1}{\epsilon_{\text{e}} n d_L^2}\right) , & \text{(E.18b)} \end{cases}$$

and

$$\epsilon_{\text{worst}}(SU(d_L)\text{-covariant code}) \geqslant \frac{1}{2n}\frac{1}{\max_i \ln d_i} + O\left(\frac{1}{n\, d_L}\right) . \tag{E.19}$$

These corollaries employ the following standard inequality of binomial coefficients. For integers $a, b > 0$, we have the two lower bounds

$$\binom{a+b}{a} \geqslant \begin{cases} \left(1 + \dfrac{a}{b}\right)^b & \text{(E.20a)} \\[2ex] \left(1 + \dfrac{b}{a}\right)^a , & \text{(E.20b)} \end{cases}$$

noting that $\binom{a+b}{b} = \binom{a+b}{a}$.

Proof of Corollary 19. Starting from (E.9), we set $a = d_L - 1$ and $b = \lceil (2n\epsilon_{\mathrm{worst}})^{-1} \rceil$ such that (E.20a) gives

$$\max_i \ln(d_i) \geqslant b \ln\left(1 + \frac{d_L - 1}{b}\right) \geqslant b \ln \frac{d_L - 1}{b} \geqslant \frac{1}{2n\epsilon_{\mathrm{worst}}} \ln\left[\frac{d_L - 1}{(2n\epsilon_{\mathrm{worst}})^{-1} + 1}\right] , \tag{E.21}$$

and hence

$$\max_i \ln(d_i) \geqslant \frac{1}{2n\epsilon_{\mathrm{worst}}} \ln(d_L - 1) - \frac{\ln\left(1 + (2n\epsilon_{\mathrm{worst}})^{-1}\right)}{2n\epsilon_{\mathrm{worst}}} . \tag{E.22}$$

This proves (E.17). ■

Proof of Corollary 20. Applying (E.20b) to (E.9), we obtain

$$\begin{aligned}\max_i d_i &\geqslant \left[\frac{d_L - 1 + \lceil (2n\epsilon_{\mathrm{worst}})^{-1} \rceil}{d_L - 1}\right]^{d_L - 1} \\ &= \exp\left\{(d_L - 1) \ln\left(1 + \frac{\lceil (2n\epsilon_{\mathrm{worst}})^{-1} \rceil}{d_L - 1}\right)\right\} .\end{aligned} \tag{E.23}$$

The bound (E.18a) follows from (E.23) by noting that $\lceil (2n\epsilon_{\mathrm{worst}})^{-1} \rceil \geqslant (2n\epsilon_{\mathrm{worst}})^{-1}$ and that $\log(1 + x) \geqslant \log(x)$. The alternative bound (E.18b) follows from the use of the bound (E.10) instead of (E.9), following the same steps while effecting the replacement $\epsilon_{\mathrm{worst}} \to d_L \epsilon_{\mathrm{e}}/2$.

Now we prove (E.19). We can rearrange (E.23) into

$$\exp\left\{\frac{\max_i \ln(d_i)}{d_L - 1}\right\} - 1 \geqslant \frac{\lceil (2n\epsilon_{\mathrm{worst}})^{-1} \rceil}{d_L - 1} \geqslant \frac{(2n\epsilon_{\mathrm{worst}})^{-1}}{d_L - 1} , \tag{E.24}$$

which in turn implies

$$\epsilon_{\mathrm{worst}} \geqslant \frac{1}{2n(d_L - 1)} \left[\max_i \left(\exp\left\{\frac{\ln(d_i)}{d_L - 1}\right\} - 1\right)\right]^{-1} . \tag{E.25}$$

Henceforth we let i denote the index of the physical subsystem with largest dimension, i.e., $d_i = \max_{i'} d_{i'}$. For large d_L, we have

$$(d_L - 1)\left(\exp\left\{\frac{\ln(d_i)}{d_L - 1}\right\} - 1\right) = \ln(d_i) + O\left(\frac{\ln^2(d_i)}{d_L}\right) = \ln(d_i)\left[1 + O\left(\frac{\ln(d_i)}{d_L}\right)\right] , \tag{E.26}$$

and thus

$$\epsilon_{\mathrm{worst}} \geqslant \frac{1}{2n \max_i \ln(d_i)}\left[1 + O\left(\frac{\ln(d_i)}{d_L}\right)\right] = \frac{1}{2n \max_i \ln(d_i)} + O\left(\frac{1}{n d_L}\right) , \tag{E.27}$$

which is the desired bound (E.19). ■

SM. F: Circumventing the Eastin-Knill theorem by randomized constructions

The proof of Theorem VII is technical, and relies on the recent developments in the representation theory of $U(d)$, and new counting formulas for the *Littlewood-Richardson* coefficients. In the first part of this section we sketch our technical argument and we provide the complete proof in the second part of the section.

F.1. Randomized constructions: Overview

Although our randomized constructions do not properly work for producing good $U(2)$-covariant codes,[8] for the $U(3)$ case we can find explicit (non-asymptotic) bounds with a slightly different scaling. There, one can benefit from the fact that the fusion rules of $U(3)$ representation theory are known [61, Section 5]. We will not discuss $U(3)$ case further, and will focus on $d_L \geqslant 4$ for the rest of this section.

Consider codes that map logical information on the Hilbert space $\mathscr{H}_L$ to three physical subsystems $\mathscr{H}_A = \mathscr{H}_{A_1} \otimes \mathscr{H}_{A_2} \otimes \mathscr{H}_{A_3}$, and denote by d_i the dimension of $\mathscr{H}_{A_i}$. In order to precisely define what we mean by the random isometry V, consider the state corresponding to V (similar to what we did in the analysis of correlation functions in SM. B). The corresponding state, $|\Psi\rangle$, lives on $\mathscr{H}_R \otimes \mathscr{H}_{A_1} \otimes \mathscr{H}_{A_2} \otimes \mathscr{H}_{A_3}$, where as before $\mathscr{H}_R \simeq \mathscr{H}_L$ is a reference system. The covariance of V translates to the invariance of $|\Psi\rangle$:

$$\left[\overline{U} \otimes r_1(U) \otimes r_2(U) \otimes r_3(U)\right] |\Psi\rangle_{RA_1A_2A_3} = |\Psi\rangle_{RA_1A_2A_3} \text{ for all } U \in U(d_L). \tag{F.1}$$

Therefore, $\Psi_{RA_1A_2A_3}$ lives on an invariant subspace of the unitary group. The projector to this invariant subspace is given by

$$\Pi_{RA_1A_2A_3} = \int dU\, \overline{U} \otimes r_1(U) \otimes r_2(U) \otimes r_3(U)\ . \tag{F.2}$$

We denote by $d_P = \mathrm{tr}\big(\Pi_{RA_1A_2A_3}\big)$ the dimension of the invariant subspace. Further, define $\Pi_{RA_i} := \mathrm{tr}_{A\setminus A_i}\big(\Pi_{RA_1A_2A_3}\big)$ and $\Pi_{\widehat{RA_i}} := \mathrm{tr}_{RA_i}\big(\Pi_{RA_1A_2A_3}\big)$.

Now, we can chose the state $|\Psi\rangle_{RA_1A_2A_3}$ randomly from $\Pi_{RA_1A_2A_3}$, and *define* V to be the corresponding isometry, i.e., $V_{L\to A_1A_2A_3} := \langle\Phi|_{LR}|\Psi\rangle_{RA_1A_2A_3}$, where $|\Phi\rangle = \sum |k\rangle_L|k\rangle_R$ for some standard choice of bases on $\mathscr{H}_L$ and $\mathscr{H}_R$.

As in SM. B, we consider single erasures at known locations, i.e., the noise channel is given by $\mathcal{N}(\cdot) = \sum q_i |i\rangle\langle i|_C \otimes \mathcal{N}^i(\cdot)$, where $\mathcal{N}^i$ erases the i-th system as per (A.1). If the isometry V is chosen at random in the space of covariant isometries, then on average, the fidelity of recovery of the code defined by the isometry is lower bounded as follows.

Lemma 21. *Suppose that the covariant isometry V is chosen randomly as above. Then, the infidelity of the code after erasure of subsystem $i \in \{1, 2, 3\}$, averaged over all covariant*

[8] More precisely, our techniques do not lead to proper lower bounds for the fidelity of recovery of random $U(2)$-covariant codes, but this might only be caused by not lower bounding the fidelity of recovery with strong enough inequalities.

isometries, satisfies the following inequality:

$$\frac{1}{2}\,\mathbb{E}[\epsilon_{\mathrm{e}}^2(\mathcal{N}^i \circ \mathcal{E})] \leqslant \frac{1}{2}\left\|\frac{\Pi_{RA_i}}{d_P} - \frac{\mathbb{1}_{RA_i}}{d_L d_i}\right\|_1 + \frac{1}{2}\sqrt{d_L d_i}\sqrt{\frac{\operatorname{tr}\left(\Pi^2_{\widehat{RA_i}}\right)}{d_P^2}}\ . \tag{F.3}$$

(Proof on page 39.)

Intuitively, Lemma 21 states that in order to get good quantum codes we need to do the followings:

1. Control the constant offset, $\left\|d_P^{-1}\Pi_{RA_i} - \mathbb{1}_{RA_i}/(d_R d_i)\right\|_1$. This can be achieved by making sure that Π_{RA_i} is close to a multiple of identity.

2. Control the *fluctuations* by minimizing $d_p^{-2}\operatorname{tr}\left(\Pi^2_{\widehat{RA_i}}\right)$. Note that this is the purity of density matrix $\Pi_{\widehat{RA_i}}/d_p$, so it would be small if $\Pi_{\widehat{RA_i}}$ is close to a multiple of a projector.

Lemma 21 is how far we can go without discussing the detailed representation theory of $U(d_L)$. From now on, we focus on analyzing Π_{RA_i} and $\Pi_{\widehat{RA_i}}$.

Without loss of generality assume that $i = 1$. Also, suppose λ, μ, ν are the Young diagrams defining the irreducible representations r_1, r_2 and r_3. Similarly, $r_{e_1}(U) = U$, where $e_1 = (1, 0, 0, \cdots, 0)$ is the Young diagram of the standard representation. Now, we use representation theory techniques to explicitly compute Π_{RA_1} and $\Pi_{\widehat{RA_1}} = \Pi_{A_2A_3}$.

The degeneracies of fusion of different irreps of $U(d_L)$ are known, and specified by the so called Littlewood-Richardson coefficients $c^{\theta}_{\mu\nu}$:

$$r_\mu \otimes r_\nu = \bigoplus_\theta r_\theta \otimes I_{c^{\theta}_{\mu\nu}}\ . \tag{F.4}$$

A specific case of this formula which is also applicable to our analysis is a version *Pieri formula* (See Appendix A.1 of [106]): If e_i is the i-th computation basis vector, then

$$\overline{r}_{e_1} \otimes r_\lambda = \bigoplus_{i\in\mathcal{I}} r_{\lambda - e_i} \tag{F.5}$$

where $\mathcal{I} \in \{1, 2, \cdots, d_L\}$ is the index set that $\lambda - e_i$ is a valid Young diagram, i.e., a non-increasing sequence. In particular, if λ is strictly decreasing then $\mathcal{I} = \{1, 2, \cdots, d_L\}$. This relation can be derived by either directly applying the Littlewood-Richardson rule [106, Appendix A.1], or starting from the standard Pieri formula and dualizing representations. With this, we have

$$\Pi_{RA_1A_2A_3} = \int dU(\bigoplus_{i\in\mathcal{I}} r_{\lambda-e_i}(U)) \otimes (\bigoplus_\theta r_\theta(U) \otimes I_{c^{\theta}_{\mu\nu}}). \tag{F.6}$$

From the Schur orthogonality relations for compact groups (Peter-Weyl theorem), we have that $\int dU[\operatorname{tr}\overline{r_\beta(U)}]r_\alpha(U) = \delta_{\alpha\beta}I_\alpha/d_\alpha$. Applying this to Eq. (F.6) we get the following explicit relations:

$$\Pi_{RA_1} = \bigoplus_{i\in\mathcal{I}} \frac{c_{\mu\nu(\lambda-e_i)}}{d_{\lambda-e_i}} I_{d_{\lambda-e_i}}, \tag{F.7}$$

$$\Pi_{A_2A_3} = \bigoplus_{i\in\mathcal{I}} \frac{1}{d_{\lambda-e_i}} I_{d_{\lambda-e_i}} \otimes I_{c_{\mu\nu(\lambda-e_i)}} , \tag{F.8}$$

where $c_{\mu\nu\lambda} := c^{\overline{\lambda}}_{\mu\nu}$, and $\overline{\lambda}$ is the dual of λ. Recall that in order for the random codes to perform well, we need that Π_{RA_1} and $\Pi_{A_2A_3}$ to be close to multiples of projectors. Equations (F.7) and (F.8) show that to achieve this we only need $\frac{c_{\mu\nu(\lambda-e_i)}}{d_{\lambda-e_i}}$ and $\frac{1}{d_{\lambda-e_i}}$ to be almost constants as i varies. The following lemma makes this observation quantitative:

Lemma 22. *Suppose that $0 \leqslant \delta \leqslant 1/2$ and c are real numbers such that for all $i \in \mathcal{I}$,*

$$1-\delta \leqslant \frac{c_{\mu\nu(\lambda-e_i)}}{c} \leqslant 1+\delta \tag{F.9}$$

$$1-\delta \leqslant \frac{d_{\lambda-e_i}}{d_\lambda} \leqslant 1+\delta, \tag{F.10}$$

then,

$$\frac{1}{2}\,\mathbb{E}[\epsilon_{\mathrm{e}}^2(\mathcal{N}^1\circ\mathcal{E})] \leqslant 4\delta + \frac{5}{2\sqrt{c}}. \tag{F.11}$$

(Proof on page 41.)

Lemma 22 demonstrates that in order to get useful lower bounds on the fidelity of the codes, one has to show that d_λ and $c_{\mu\nu\lambda}$ are stable under perturbations by e_i. We construct our irreps such that they achieve this stability.

Define $|\lambda| := \sum_i \lambda_i$ for arbitrary Young diagram λ. It is known that if $|\mu|+|\nu|+|\lambda| \neq 0$, then $c_{\mu\nu\lambda} = 0$. Now, the construction is as follows: Fix a triplet of Young diagrams $(\hat\mu, \hat\nu, \hat\lambda)$ such that $|\hat\mu_i| + |\hat\nu_i| + |\hat\lambda_i| = 0$ and set

$$(\mu,\nu,\lambda) = (N\hat\mu + e_1, N\hat\nu, N\hat\lambda) , \tag{F.12}$$

for large values of N. We used $N\hat\mu+e_1$ instead of $N\hat\mu$ is to ensure that $|\mu|+|\nu|+|\lambda-e_i| = 0$ as we need $c_{\mu\nu(\lambda-e_i)}$ to be non-zero.

Showing smoothness of d_λ is much simpler, because by the *Weyl dimension formula* (see [106, Section 15.3]) it is polynomial in $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{d_L})$. So by basic Taylor expansion we have $d_{\lambda+e_i} = d_\lambda + \partial d_\lambda/\partial\lambda_i + 1/2\,\partial^2 d_\lambda/\partial\lambda_i^2 + \cdots$. Note that the total degree of the terms in the sum decreases by differentiation. Hence, d_λ is the dominant term in the expansion of $d_{\lambda+e_i}$ and other terms are lower order in N. This is true when the $\hat\lambda_i$'s are non-zero, and different from each other, so the the first term in the Taylor expansion is non-zero. Hence, in this proof, we always assume that $\hat\mu$ and $\hat\nu$, and $\hat\lambda$ have rows different from each other and zero. This condition will naturally hold with high probability if the diagrams are chosen randomly. Therefore, there exist N_0' and C_0' such that for $N \geqslant N_0'$,

$$1 - \frac{C_0'}{N} \leqslant \frac{d_{\lambda-e_i}}{d_\lambda} \leqslant 1 + \frac{C_0'}{N}. \tag{F.13}$$

The Littlewood-Richardson coefficients are much more complicated. They can be computed using efficient algorithms, such as the Littlewood-Richardson rule, but there is no explicit formula. In fact, they are specific cases of the called *Kronecker coefficients* whose computation is known to be NP-hard [107]. However, a series of new developments in the representation theory of the unitary group has revealed interesting polynomiality properties for the LR coefficients.

It is known that $c_{\mu\nu\lambda}$ as a function of μ, ν and λ (in the $3d_L - 1$ dimensional subspace constrained by the condition $|\mu| + |\nu| + |\lambda| = 0$) is non-zero if and only if (μ, ν, λ) is in a particular convex cone. This cone, or *chamber complex*, is then divided to several sub-cones or *chambers*. In Ref. [61] it is shown that $c_{\mu\nu\lambda}$ is a polynomial within each chamber (see Fig. 5 of the main text).

We choose $(\hat{\mu}, \hat{\nu}, \hat{\lambda})$ such that it is in the interior of one of the chambers, and $c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}$ is not constant. If N is large enough, $c_{\mu\nu(\lambda-e_i)}$'s will remain in the interior of the same chamber for all i, and are described by the same polynomial. Therefore, similar to d_λ,[9] we have:

$$1 - \frac{C_0''}{N} \leqslant \frac{c_{\mu\nu(\lambda-e_i)}}{c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}} \leqslant 1 + \frac{C_0''}{N} \ , \tag{F.14}$$

where $N \geqslant N_0''$ and for some N_0'' and C_0'', and $c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}$ plays the role of c in Lemma 22. Clearly, this bounds show the smoothness required for Lemma 22 to work, and therefore we get our main theorem:

Theorem 23. *Suppose that $d_L \geqslant 4$. There exist Young diagrams $\hat{\lambda}$, $\hat{\mu}$ and $\hat{\nu}$, an integer N_0, and a constant C_0, such that if $(\mu, \nu, \lambda) = (N\hat{\mu} + e_1, N\hat{\nu}, N\hat{\lambda})$ and V is a random covariant isometry in the sense of (57), we have,*

$$\epsilon_{\mathrm{e}}(\mathcal{N} \circ \mathcal{E}) \leqslant \frac{C_0}{\sqrt{N}} \qquad \textit{for} \qquad N \geqslant N_0 \ . \tag{F.15}$$

For these constructions, we have,

$$\epsilon_{\mathrm{e}}(\mathcal{N} \circ \mathcal{E}) \leqslant C_1 \, (\max_i d_i)^{-\frac{1}{d_L(d_L-1)}} \ . \tag{F.16}$$

(Proof on page 42.)

Finally, Theorem VII follows immediately from Theorem 23.

F.2. Randomized constructions: Detailed proofs

First, we prove Lemma 21.

Proof of Lemma 21. First, we express the error-correcting accuracy of the code V according to the average entanglement fidelity in terms of the distance of the codewords to a maximally mixed state, including the reference system. By Bény/Oreshkov (12a) (choosing $\zeta_E = \mathbb{1}_{A_i}/d_i$), we have

$$f_{\mathrm{e}}(\mathcal{N}^i \circ \mathcal{E}) \geqslant F\big(\widehat{\mathcal{N}^i \circ \mathcal{E}}(\hat{\phi}_{LR}), \zeta_E \otimes \hat{\phi}_R\big) = F\left(\Psi_{RA_i}, \frac{\mathbb{1}_{A_i}}{d_i} \otimes \frac{\mathbb{1}_R}{d_L}\right) \ , \tag{F.17}$$

[9] See the discussion below Eq. (F.39) as well.

and hence

$$\epsilon_{\mathrm{e}}^2(\mathcal{N}^i\circ\mathcal{E}) \leqslant 1 - F^2\big(\Psi_{RA_i}, \frac{\mathbb{1}_{RA_i}}{d_L d_i}\big) \leqslant \left\|\Psi_{RA_i} - \frac{\mathbb{1}_{RA_i}}{d_L d_i}\right\|_1 , \tag{F.18}$$

where we recall the usual relations between trace distance and the fidelity.

We denote by $\mathbb{E}$ the averaging over all possible invariant states $\Psi_{RA_1A_2A_3}$. Taking an average over (F.18) gives us

$$\frac{1}{2}\,\mathbb{E}(\epsilon_{\mathrm{e}}^2(\mathcal{N}^i\circ\mathcal{E})) \leqslant \frac{1}{2}\,\mathbb{E}\,\|\Psi_{RA_i} - \tau_{RA_i}\|_1 , \tag{F.19}$$

where we write $\tau_{RA_i} = \mathbb{1}_{RA_i}/(d_L d_i)$. Applying triangle inequality, Cauchy-Schwarz inequality, and the concavity of square root gives us (see Ref. [32] for similar calculations),

$$\begin{aligned}\frac{1}{2}\,\mathbb{E}(\epsilon_{\mathrm{e}}^2(\mathcal{N}^i\circ\mathcal{E})) &\leqslant \frac{1}{2}\,\mathbb{E}\,\|\Psi_{RA_i} - \tau_{RA_i}\|_1\\ &\leqslant \frac{1}{2}\|\mathbb{E}\,\Psi_{RA_i} - \tau_{RA_i}\|_1 + \frac{1}{2}\,\mathbb{E}\|\Psi_{RA_i} - \mathbb{E}\,\Psi_{RA_i}\|_1\\ &\leqslant \frac{1}{2}\|\mathbb{E}\,\Psi_{RA_i} - \tau_{RA_i}\|_1 + \frac{1}{2}\sqrt{d_R d_i\left(\mathrm{tr}[\mathbb{E}\,\Psi_{RA_i}^2] - \mathrm{tr}[(\mathbb{E}\,\Psi_{RA_i})^2]\right)}.\end{aligned} \tag{F.20}$$

Now, consider the rank-d_P projector to the invariant space $\Pi_{RA_1A_2A_3}$ that we constructed in §VII B. Define $L:\mathbb{C}^{d_P}\to\mathscr{H}_A\otimes\mathscr{H}_R$ be the isometry mapping to the invariant space, which satisfies

$$L^\dagger L = \mathbb{1}_{d_P}\ ;\ \text{and} \qquad\qquad LL^\dagger = \Pi_{RA_1A_2A_3}\ . \tag{F.21}$$

We can define $|\Psi\rangle_{RA} = L|\chi\rangle$, where $|\chi\rangle$ is a random state in $\mathbb{C}^{d_P}$. Then,

$$\mathbb{E}\,\Psi_{RA_i} = \mathbb{E}\,\mathrm{tr}_{AA_i}\big(L\chi L^\dagger\big) = \mathrm{tr}_{\widehat{RA_i}}\big(L(\mathbb{E}\,\chi)L^\dagger\big) = \frac{1}{d_P}\,\mathrm{tr}_{\widehat{RA_i}}\big(LL^\dagger\big) = \frac{\Pi_{RA_i}}{d_P}\ , \tag{F.22}$$

where we used $\mathbb{E}\,\chi = \mathbb{1}/d_P$. For simplicity, we henceforth set $i=1$ without loss of generality. If $\mathcal{F}_{A_2A_3}$ is flip operator swapping two copies of the Hilbert space $\mathscr{H}_{A_2A_3} = \mathscr{H}_{A_2}\otimes\mathscr{H}_{A_3}$, we have

$$\begin{aligned}\mathbb{E}\,\mathrm{tr}[\Psi_{RA_1}^2] &= \mathbb{E}\,\mathrm{tr}[\Psi^{\otimes 2}\mathcal{F}_{A_2A_3}] = \mathrm{tr}[L^{\otimes 2}\,\mathbb{E}\,\chi^{\otimes 2}L^{\dagger\otimes 2}\mathcal{F}_{A_2A_3}]\\ &= \mathrm{tr}[L^{\otimes 2}\frac{I+\mathcal{F}}{d_P(d_P+1)}L^{\dagger\otimes 2}\mathcal{F}_{A_2A_3}] = \frac{1}{d_P(d_P+1)}\,\mathrm{tr}[\Pi_{RA}^{\otimes 2}(\mathcal{F}_{RA_1}+\mathcal{F}_{A_2A_3})]\\ &= \frac{\mathrm{tr}(\Pi_{RA_1}^2)+\mathrm{tr}(\Pi_{A_2A_3}^2)}{d_P(d_P+1)}\ .\end{aligned} \tag{F.23}$$

Substituting into (F.20) and applying basic inequalities lead to,

$$\begin{aligned}\frac{1}{2}\,\mathbb{E}(\epsilon_{\mathrm{e}}^2(\mathcal{N}^1\circ\mathcal{E})) &\leqslant \frac{1}{2}\left\|\frac{\Pi_{RA_i}}{d_P} - \tau_{RA_1}\right\|_1 + \frac{1}{2}\frac{\sqrt{d_R d_1}}{d_P}\sqrt{\frac{1}{1+1/d_P}\,\mathrm{tr}(\Pi_{A_2A_3}^2) - \frac{1}{d_P+1}\,\mathrm{tr}(\Pi_{RA_1}^2)}\\ &\leqslant \frac{1}{2}\left\|\frac{\Pi_{RA_1}}{d_P} - \tau_{RA_1}\right\|_1 + \frac{1}{2}\frac{\sqrt{d_R d_1}}{d_P}\sqrt{\mathrm{tr}(\Pi_{A_2A_3}^2)}\ ,\end{aligned} \tag{F.24}$$

which is the desired formula. ■

Next, we prove Lemma 22.

Proof of Lemma 22. For simplicity of exposition, define two probability distributions $p, q : \mathcal{I} \to \mathbb{R}_{\geqslant 0}$,

$$p_i = \frac{c_{\mu\nu(\lambda-e_i)}}{d_P} \ ; \qquad q_i = \frac{d_{\lambda-e_i}}{d_R d_1} \ . \tag{F.25}$$

From Lemma 21, we have,

$$\frac{1}{2}\,\mathbb{E}(\epsilon_{\mathrm{e}}^2(\mathcal{N}^i \circ \mathcal{E})) \leqslant \frac{1}{2}\left\| \frac{\Pi_{RA_i}}{d_P} - \tau_{RA_i} \right\|_1 + \frac{1}{2}\sqrt{d_L d_1}\sqrt{\frac{\operatorname{tr}\Pi^2_{A_2A_3}}{d_P^2}} \ . \tag{F.26}$$

We would like to bound both terms on the right hand side of (F.26). We have

$$\left\| \frac{\Pi_{RA_i}}{d_P} - \tau_{RA_i} \right\|_1 = \sum_{i\in\mathcal{I}} d_{\lambda-e_i}\left| \frac{c_{\mu\nu(\lambda-e_i)}}{d_P d_{\lambda-e_i}} - \frac{1}{d_R d_1} \right| = \sum_{i\in\mathcal{I}} |p(i) - q(i)| \ . \tag{F.27}$$

Also,

$$\operatorname{tr}\bigl(\Pi^2_{A_2A_3}\bigr) = \sum_{i\in\mathcal{I}} (d_{\lambda-e_i} c_{\mu\nu(\lambda-e_i)})\frac{1}{d^2_{\lambda-e_i}} = \frac{d_P}{d_R d_1}\sum_{i\in\mathcal{I}}\frac{p(i)}{q(i)} \ . \tag{F.28}$$

Now, the condition of the lemma can be written as

$$1-\delta \leqslant \frac{p_i}{c/d_P} \leqslant 1+\delta \ . \tag{F.29}$$

By summing over i, we get,

$$\frac{1}{|\mathcal{I}|(1+\delta)} \leqslant \frac{c}{d_P} \leqslant \frac{1}{|\mathcal{I}|(1-\delta)} \ . \tag{F.30}$$

With some algebra, we obtain

$$\left| p_i - \frac{1}{|\mathcal{I}|} \right| \leqslant \left| p_i - \frac{c}{d_P} \right| + \left| \frac{c}{d_P} - \frac{1}{|\mathcal{I}|} \right| \leqslant \frac{\delta}{|\mathcal{I}|(1-\delta)} + \frac{\delta}{|\mathcal{I}|(1-\delta)} = \frac{4\delta}{|\mathcal{I}|} \ . \tag{F.31}$$

Similarly, $\left| q_i - 1/|\mathcal{I}| \right| \leqslant 4\delta/|\mathcal{I}|$. Therefore,

$$\left\| \frac{\Pi_{RA_i}}{d_P} - \tau_{RA_i} \right\|_1 = \sum_{i\in\mathcal{I}} |p_i - q_i| \leqslant \sum_{i\in\mathcal{I}} \left| p_i - \frac{1}{|\mathcal{I}|} \right| + \left| q_i - \frac{1}{|\mathcal{I}|} \right| \leqslant 8\delta \ . \tag{F.32}$$

On the other hand, $p_i \leqslant (1+\delta)\frac{c}{d_P}$, and $1/q_i \leqslant \frac{d_R d_1}{d_\lambda(1-\delta)}$. Now, we get that $p_i/q_i \leqslant (1+\delta)^2/(1-\delta)^2$. So,

$$\operatorname{tr}\bigl(\Pi^2_{A_2A_3}\bigr) = \frac{d_P}{d_R d_1}\sum_{i\in\mathcal{I}}\frac{p_i}{q_i} \leqslant \frac{d_P|\mathcal{I}|}{d_R d_1}\left(\frac{1+\delta}{1-\delta}\right)^2$$

$$\leqslant \frac{d_P^2}{d_R d_1 c}\frac{(1+\delta)^2}{(1-\delta)^3} \leqslant 5^2 \frac{d_P^2}{d_R d_1 c} \ . \tag{F.33}$$

Substituting in the formula for the fidelity completes the proof. ∎

Next, we would like to prove our main theorem on random constructions, Theorem 23. Before that, we need to show that the Littlewood-Richardson coefficients can grow significantly with the size the Young diagrams. This is the content of next lemma:

Lemma 24. *In the chamber complex of Littlewood-Richardson coefficients discussed in §VII B, there are chambers in which $c_{\mu\nu\lambda}$ is a polynomial of degree $\binom{d_L-1}{2}$*

Proof of Lemma 24. Consider the following relation for the Littlewood-Richardson coefficients, derived by comparing dimensions:

$$d_\mu d_\nu = \sum_\lambda c_{\mu\nu\lambda} d_\lambda. \tag{F.34}$$

Define the average of Littlewood-Richardson coefficients weighted by the dimension d_λ, i.e.,

$$\overline{c} = \frac{\sum_\lambda c_{\mu\nu\lambda} d_\lambda}{\sum_\lambda d_\lambda} \ . \tag{F.35}$$

Also, assume that the number of λ's where $c_{\mu\nu\lambda} \neq 0$ is $N_{\mu\nu}$ and the average dimension of d_λ, averaged over such λ's is,

$$\overline{d} = \frac{\sum_{\lambda \text{ where } c_{\mu\nu\lambda}\neq 0} d_\lambda}{N_{\mu\nu}} \ . \tag{F.36}$$

Now (F.34) becomes

$$\frac{d_\mu d_\nu}{N_{\mu\nu}\overline{d}} = \overline{c} \ . \tag{F.37}$$

Consider the case where $\mu = N\mu_0$ and $\nu = N\nu_0$, for some fixed μ_0 and ν_0 and large N. It is known that the dimension of the chamber complex is $3d_L - 1$, see, e.g., Proposition 1 in [108]. Therefore, as two d_L dimensional axis are fixed by μ and ν, the section of the cone corresponding to $c_{\mu\nu\lambda} \neq 0$ is $d_L - 1$ dimensional, and therefore $N_{\mu\nu} = O(N^{d_L-1})$. From the Weyl dimension formula, it is known that d_μ, d_ν , and $\overline{d}$ are all $O\left(N^{d_L(d_L-1)/2}\right)$. So,

$$\overline{c} = O\left((N^{(d_L-1)(d_L-2)/2}\right) \ . \tag{F.38}$$

This shows that there exists at least one chamber whose polynomial is at least degree $\binom{d_L-1}{2}$. On the other hand, it is known that degree of the polynomials are bounded above by $\binom{d_L-1}{2}$ (see Corollary 4.2 in [61]). This completes the proof. ∎

Proof of Theorem 23. We start from Eqs. (F.13) and (F.14). If we set $C_0 = \max(C_0', C_0'')$ and $N_0 = \max(N_0', N_0'')$, we have

$$1 - \frac{C_0}{N} \leqslant \quad \tfrac{d_{\lambda-e_i}}{d_\lambda} \quad \leqslant 1 + \frac{C_0}{N} \ ; \tag{F.39a}$$

$$1 - \frac{C_0}{N} \leqslant \frac{c_{\mu\nu(\lambda-e_i)}}{c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}} \leqslant 1 + \frac{C_0}{N} \ . \tag{F.39b}$$

If one is unlucky, it is possible to choose points in the chamber complex such that the degree of $c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}$ as a polynomial in N is less than the degree of the multivariate polynomial $c_{\mu,\nu,\lambda}$ as a polynomial in $\{\lambda_i, \nu_j, \mu_k\}$ in the corresponding chamber. This can happen only when highest degree terms in N vanish in $c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}$, however, this vanishing happens on a measure-zero set of points, and can be avoided. Further, suppose that $\hat{\mu}$,$\hat{\nu}$, and $\hat{\lambda}$ were chosen such that $c_{N\hat{\mu},N\hat{\nu},N\hat{\lambda}}$ grows superlinearly as a function of N. This is possible for $d_L \geqslant 4$ as a result of Lemma 24. Using this fact and Lemma 22, we get

$$\mathbb{E}(\epsilon_{\mathrm{e}}^2(\mathcal{N}^1 \circ \mathcal{E})) = O(1/N) \ . \tag{F.40}$$

In fact, the same relation holds for $\epsilon_{\mathrm{e}}(\mathcal{N}^2 \circ \mathcal{E})$ and $\epsilon_{\mathrm{e}}(\mathcal{N}^3 \circ \mathcal{E})$, and using the Markov inequality and the union bound we can show that there exists $\hat{\mu}$,$\hat{\nu}$, and $\hat{\lambda}$ for which

$$\max\bigl(\epsilon_{\mathrm{e}}^2(\mathcal{N}^1 \circ \mathcal{E}), \epsilon_{\mathrm{e}}^2(\mathcal{N}^2 \circ \mathcal{E}), \epsilon_{\mathrm{e}}^2(\mathcal{N}^3 \circ \mathcal{E})\bigr) = O(1/N) \ . \tag{F.41}$$

As a consequence, and using Lemma 29, we get (F.15). The second equation, (F.16), follows from (F.15) using the Weyl dimension which indicates that $d_i = O\left(N^{d_L(d_L-1)/2}\right)$. ∎

SM. G: General lemmas

We begin with a simple characterization of the quantity ΔA.

Proposition 25. *For any Hermitian A, we have*

$$\Delta A = 2 \left\| A - \frac{a_+ + a_-}{2} \mathbb{1} \right\|_\infty = 2 \min_{\nu \in \mathbb{R}} \| A - \nu \mathbb{1} \|_\infty \ , \tag{G.1}$$

where $a_\pm \in \mathbb{R}$ are the maximum and minimum eigenvalues of A (such that $\Delta A = a_+ - a_-$).

Proof of Proposition 25. Define $A' = A - [(a_+ + a_-)/2]\mathbb{1}$ and observe that $\Delta A' = \Delta A$. Because the minimum and maximum eigenvalues of A' differ only by a sign, we have $\|A'\|_\infty = \Delta A'/2$. For the second equality, let a' be the maximum eigenvalue of A', noting that $-a'$ is therefore the corresponding minimum eigenvalue. Observe that for any $\nu' \in \mathbb{R}$, we have $\|A' + \nu'\mathbb{1}\|_\infty = \max(a' + \nu', -(-a' + \nu'))$ because the infinity norm of a Hermitian matrix is either the maximum eigenvalue or the negative minimum eigenvalue of its argument, whichever is greater. Hence $2\|A' + \nu'\mathbb{1}\|_\infty = 2\max(a' + \nu', a' - \nu') = 2a' + 2|\nu'| \geqslant 2a' = 2\|A'\|_\infty = \Delta A' = \Delta A$. Therefore, for any $\nu \in \mathbb{R}$, $2\|A - \nu\mathbb{1}\|_\infty \geqslant 2\|A'\|_\infty$, proving the second equality. ∎

The following semidefinite problem appears multiple times across this supplemental material, and is therefore formulated in a separate proposition:

Proposition 26. *For any Hermitian operator T and any positive semidefinite operator ξ,*

$$\min_{\mu \in \mathbb{R}} \left\| \xi^{1/2}(T - \mu\mathbb{1})\xi^{1/2} \right\|_1 = \max_{\substack{\|X\|_\infty \leqslant 1 \\ \mathrm{tr}(X\xi)=0}} \mathrm{tr}\bigl(\xi^{1/2} T \xi^{1/2} X\bigr) \ , \tag{G.2}$$

where the maximization is taken over all Hermitian operators X satisfying the given con-

straints.

Proof of Proposition 26. The optimization on the right hand side of (G.2) is a semidefinite program, and we proceed to compute its dual program [109]. In terms of the variables $X = X^\dagger$, $A, B \geqslant 0$, and $\mu \in \mathbb{R}$, we have

$$\begin{aligned}
&\max_{\substack{\|X\|_\infty \leqslant 1 \\ \mathrm{tr}(X\xi)=0}} \mathrm{tr}\big[\xi^{1/2}\, T\, \xi^{1/2}\, X\big] \\
&\quad = \quad \begin{array}{rll} \text{maximize}: & \mathrm{tr}\big[\xi^{1/2}\, T\, \xi^{1/2}\, X\big] & \\ \text{over variable}: & X = X^\dagger & \\ \text{subject to}: & X \leqslant \mathbb{1}_L & : A \\ & X \geqslant -\mathbb{1}_L & : B \\ & \mathrm{tr}(X\xi) = 0 & : \mu \end{array} \qquad \text{(G.3a)} \\
&\quad = \quad \begin{array}{rl} \text{minimize}: & \mathrm{tr}(A) + \mathrm{tr}(B) \\ \text{over variables}: & A \geqslant 0; \quad B \geqslant 0; \quad \mu \in \mathbb{R} \\ \text{subject to}: & \xi^{1/2} T \xi^{1/2} = \mu\xi + A - B\ . \end{array} \qquad \text{(G.3b)}
\end{aligned}$$

Strong duality holds because of Slater's conditions [109]. Indeed $X = 0$ is strictly feasible in the primal problem; the dual is actually also strictly feasible by choosing (say) $\mu = 0$ and A and B to be the positive and negative parts respectively of the Hermitian operator $\xi^{1/2} T \xi^{1/2}$ plus a constant times the identity. Recall that for any Hermitian operator T' we have

$$\|T'\|_1 = \min_{\substack{\Delta_\pm \geqslant 0 \\ T' = \Delta_+ - \Delta_-}} \mathrm{tr}(\Delta_+) + \mathrm{tr}(\Delta_-)\ . \tag{G.4}$$

For a fixed μ in (G.3b), we recognize the remaining optimization as the one-norm of the operator $\xi^{1/2}(T - \mu\mathbb{1})\xi^{1/2}$, and hence,

$$\text{(G.3b)} = \min_{\mu \in \mathbb{R}} \big\|\xi^{1/2}(T - \mu\mathbb{1})\xi^{1/2}\big\|_1\ , \tag{G.5}$$

proving the claim. ■

Proposition 27. *Let A be a Hermitian operator in a Hilbert space of dimension d. Then*

$$s(A) = d^{-1} \min_{x \in \mathbb{R}} \|A - x\mathbb{1}\|_1\ . \tag{G.6}$$

Proof of Proposition 27. Let μ be a median eigenvalue of A. Recall that by definition, μ is such that the length-d vector of eigenvalues of A counted with multiplicity has at least $\lceil d/2 \rceil$ components that are less than or equal to μ, and at least $\lceil d/2 \rceil$ components that are greater than or equal to μ.

Let $\{|k\rangle\}$ for $k = 1, \ldots, d$ be an eigenbasis of A with its elements arranged such that the eigenvalues of A are nonincreasing in k, $\langle 1|A|1\rangle \geqslant \langle 2|A|2\rangle \geqslant \cdots \geqslant \langle d|A|d\rangle$. Let

$$P_+ = \sum_{k=1}^{\lfloor d/2 \rfloor} |k\rangle\langle k|\ ; \qquad\qquad P_- = \sum_{k=\lceil d/2 \rceil + 1}^{d} |k\rangle\langle k|\ , \tag{G.7}$$

noting that the two projectors are orthogonal and that $\mathrm{rank}(P_+) = \mathrm{tr}(P_+) = \mathrm{tr}(P_-) = \mathrm{rank}(P_-)$. That is, we divide all basis vectors into two sets of equal size, corresponding to the smallest eigenvalues and the largest eigenvalues respectively, possibly leaving out the middle basis vector if the space dimension is odd. Then, the eigenvalues corresponding to the eigenbasis vectors included in P_+ (respectively, P_-) are all greater than or equal to (respectively less than or equal to) μ. If d is odd, then the eigenvalue associated with the basis vector that was left out is μ.

Set $X = P_+ - P_-$ which satisfies $\|X\|_\infty \leqslant 1$. We have $\|A - \mu\mathbb{1}\|_1 = \mathrm{tr}\big[X(A - \mu\mathbb{1})\big]$: Indeed, the one-norm of a Hermitian matrix is equal to the sum of the absolute values of the eigenvalues of its argument, which is precisely taken care of by our careful choice of X. Then, since $\mathrm{tr}(X) = 0$ by construction,

$$s(A) = \frac{1}{d}\,\|A - \mu\mathbb{1}\|_1 = \frac{1}{d}\,\mathrm{tr}\big[X(A - \mu\mathbb{1})\big] = \frac{1}{d}\,\mathrm{tr}\big(XA\big) \leqslant \max_{\substack{\|X\|_\infty \leqslant 1 \\ \mathrm{tr}(X)=0}} \frac{1}{d}\,\mathrm{tr}\big(XA\big)\ . \tag{G.8}$$

On the other hand we clearly have

$$s(A) = d^{-1}\|A - \mu\mathbb{1}\|_1 \geqslant d^{-1}\min_{x\in\mathbb{R}} \|A - x\mathbb{1}\|_1\ . \tag{G.9}$$

Proposition 26 (with $\xi = \mathbb{1}/d$) states that the right hand sides of Eqs. (G.8) and (G.9) are equal, proving the claim. ■

The following lemma relates the correctability of the code to the environment's ability to distinguish two states in terms of the trace distance.

Lemma 28. *For any encoding channel $\mathcal{E}$ and noise channel $\mathcal{N}$, and for any two logical states σ_L, σ'_L, and if $\widehat{\mathcal{N}\circ\mathcal{E}}$ is a complementary channel of $\mathcal{N}\circ\mathcal{E}$, we have that*

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \frac{1}{2}\,\delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_L), \widehat{\mathcal{N}\circ\mathcal{E}}(\sigma'_L)\big)\ . \tag{G.10}$$

Proof of Lemma 28. Let ζ be the state achieving the optimum in (12b). We have

$$\begin{aligned}
\epsilon^2_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) &= 1 - f^2_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \\
&= 1 - \min_{\phi_{LR}} F^2(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}), \mathcal{T}_\zeta(\phi_{LR})) \\
&= \max_{\phi_{LR}}\big[1 - F^2(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}), \zeta\otimes\phi_R)\big] \\
&\geqslant \max_{\phi_{LR}} \delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\phi_{LR}), \zeta\otimes\phi_R\big)^2\ ,
\end{aligned} \tag{G.11}$$

recalling that the trace distance obeys $\delta(\rho,\sigma) \leqslant \sqrt{1 - F^2(\rho,\sigma)}$ (see, e.g., [104]). Choosing the optimization candidates $\sigma_L \otimes |0\rangle\langle 0|_R$ and $\sigma'_L \otimes |0\rangle\langle 0|_R$ in the last inequality, we obtain both

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_L), \zeta\big)\ ; \tag{G.12}$$

$$\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E}) \geqslant \delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma'_L), \zeta\big)\ . \tag{G.13}$$

Hence, by triangle inequality,

$$\delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_L), \widehat{\mathcal{N}\circ\mathcal{E}}(\sigma'_L)\big) \leqslant \delta\big(\widehat{\mathcal{N}\circ\mathcal{E}}(\sigma_L), \zeta\big) + \delta\big(\zeta, \widehat{\mathcal{N}\circ\mathcal{E}}(\sigma'_L)\big) \leqslant 2\,\epsilon_{\mathrm{worst}}(\mathcal{N}\circ\mathcal{E})\ . \qquad \blacksquare$$

The following lemma relates the global fidelity of the code to the fidelities corresponding to the correction of individual errors. Note that we do not necessarily expect a similar result to hold for the worst-case entanglement fidelity, because the worst-case input state might be different for each erasure event.

Lemma 29. *Let $\mathcal{N}^{\alpha}_{A\to A}$ and $\mathcal{N}_{A\to AC}(\cdot) = \sum q_\alpha \mathcal{N}^\alpha(\cdot)\otimes|\alpha\rangle\langle\alpha|_C$ correspond to a noise model of erasures at known locations, as given in (A.1). Then, for any $|\phi\rangle_{LR}$, the average entanglement fidelity of the code with respect to $|\phi\rangle_{LR}$ is directly related to the individual fidelities of recovery for each possible erasure:*

$$f^2_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) = \sum q_\alpha\, f^2_{|\phi\rangle}(\mathcal{N}^\alpha\circ\mathcal{E})\ , \tag{G.14}$$

and consequently,

$$\epsilon^2_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) = \sum q_\alpha\, \epsilon^2_{|\phi\rangle}(\mathcal{N}^\alpha\circ\mathcal{E})\ . \tag{G.15}$$

Proof of Lemma 29. The average entanglement fidelity associated with the different noise channels can be written as:

$$f^2_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) = \max_{\mathcal{R}}\,\langle\phi|_{LR}\big[\mathcal{R}\circ\mathcal{N}\circ\mathcal{E}(\phi_{LR})\big]|\phi\rangle_{LR}\ ; \tag{G.16a}$$

$$f^2_{|\phi\rangle}(\mathcal{N}^\alpha\circ\mathcal{E}) = \max_{\mathcal{R}^\alpha}\,\langle\phi|_{LR}\big[\mathcal{R}^\alpha\circ\mathcal{N}^\alpha\circ\mathcal{E}(\phi_{LR})\big]|\phi\rangle_{LR}\ , \tag{G.16b}$$

where the optimizations range over recovery channels $\mathcal{R}_{AC\to L}$ and $\mathcal{R}^\alpha_{A\to L}$, respectively. We have

$$\begin{aligned}
&\max_{\mathcal{R}}\,\langle\phi|_{LR}\big[\mathcal{R}\circ\mathcal{N}\circ\mathcal{E}(\phi_{LR})\big]|\phi\rangle_{LR}\\
&\qquad = \max_{\mathcal{R}}\sum q_\alpha\langle\phi|_{LR}\big[\mathcal{R}\big(|\alpha\rangle\langle\alpha|_C\otimes\big(\mathcal{N}^\alpha\circ\mathcal{E}\big)(\phi_{LR})\big)\big]|\phi\rangle_{LR}\\
&\qquad \leqslant \sum q_\alpha \max_{\mathcal{R}}\langle\phi|_{LR}\big[\mathcal{R}\big(|\alpha\rangle\langle\alpha|_C\otimes\big(\mathcal{N}^\alpha\circ\mathcal{E}\big)(\phi_{LR})\big)\big]|\phi\rangle_{LR}\\
&\qquad \leqslant \sum q_\alpha \max_{\mathcal{R}^\alpha_{A\to L}}\langle\phi|_{LR}\big[\mathcal{R}^\alpha\big(\big(\mathcal{N}^\alpha\circ\mathcal{E}\big)(\phi_{LR})\big)\big]|\phi\rangle_{LR}\ ,
\end{aligned} \tag{G.17}$$

showing that

$$f^2_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) \leqslant \sum q_\alpha f^2_{|\phi\rangle}(\mathcal{N}^\alpha\circ\mathcal{E})\ . \tag{G.18}$$

Physically, the reverse inequality follows because a global recovery strategy is to measure the register containing the record that indicates which error occurred, and to apply the optimal recovery strategy corresponding to that error. Specifically, if $\mathcal{R}^\alpha_{A\to L}$ are optimal choices in (G.16b) for each α, then we define

$$\mathcal{R}_{AC\to L}(\cdot) = \sum \mathcal{R}^\alpha_{A\to L}\big(\langle\alpha|(\cdot)|\alpha\rangle_C\big)\ . \tag{G.19}$$

Then,

$$
\begin{aligned}
f^2_{|\phi\rangle}(\mathcal{N}\circ\mathcal{E}) &\geqslant \langle\phi|_{LR}\big[\mathcal{R}\circ\mathcal{N}\circ\mathcal{E}(\phi_{LR})\big]|\phi\rangle_{LR} \\
&= \sum q_\alpha \langle\phi|_{LR}\big[\mathcal{R}\big(|\alpha\rangle\langle\alpha|_C\otimes\big(\mathcal{N}^\alpha\circ\mathcal{E}\big)(\phi_{LR})\big)\big]|\phi\rangle_{LR} \\
&= \sum q_\alpha \langle\phi|_{LR}\big[\mathcal{R}^\alpha\big(\big(\mathcal{N}^\alpha\circ\mathcal{E}\big)(\phi_{LR})\big)\big]|\phi\rangle_{LR} \\
&= \sum q_\alpha f^2_{|\phi\rangle}(\mathcal{N}^\alpha\circ\mathcal{E}) \ ,
\end{aligned}
\tag{G.20}
$$

as claimed. ∎

The following lemma is a technical consequence of the concavity of the fidelity function.

Lemma 30. *Let ρ,σ be two (normalized) quantum states. Let $\tau\geqslant 0$ with $\rho\geqslant\tau$. Then*

$$
F(\rho,\sigma)\geqslant \mathrm{tr}(\tau)\,F\Big(\frac{\tau}{\mathrm{tr}(\tau)},\sigma\Big) \ . \tag{G.21}
$$

Proof of Lemma 30. Since $\rho\geqslant\tau$, we have $\rho-\tau=:\Delta\geqslant 0$. Then $\rho=\tau+\Delta=\mathrm{tr}(\tau)\,\frac{\tau}{\mathrm{tr}(\tau)}+\mathrm{tr}(\Delta)\,\frac{\Delta}{\mathrm{tr}(\Delta)}$, and by concavity of the fidelity,

$$
F(\rho,\sigma)=F\Big(\mathrm{tr}(\tau)\,\frac{\tau}{\mathrm{tr}(\tau)}+\mathrm{tr}(\Delta)\,\frac{\Delta}{\mathrm{tr}(\Delta)},\sigma\Big)\geqslant \mathrm{tr}(\tau)\,F\Big(\frac{\tau}{\mathrm{tr}(\tau)},\sigma\Big)+\mathrm{tr}(\Delta)\,F\Big(\frac{\Delta}{\mathrm{tr}(\Delta)},\sigma\Big) \ . \tag{G.22}
$$

The claim follows by noting that $\mathrm{tr}(\Delta)\,F\big(\Delta/\mathrm{tr}(\Delta),\sigma\big)\geqslant 0$. ∎

[1–103] See main text.

[104] M. Tomamichel, R. Colbeck, and R. Renner, "Duality between smooth min- and max-entropies," IEEE Transactions on Information Theory **56**, 4674–4681 (2010), arXiv:0907.5238.

[105] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Information Theory*, Ph.D. thesis, ETH Zurich (2012), arXiv:1203.2142.

[106] W. Fulton and J. Harris, *Representation theory: a first course*, Vol. 129 (Springer Science & Business Media, 2013).

[107] C. Ikenmeyer, K. D. Mulmuley, and M. Walter, "On vanishing of Kronecker coefficients," computational complexity **26**, 949–992 (2017), arXiv:1507.02955.

[108] A. Knutson, T. Tao, and C. Woodward, "The honeycomb model of $gl_n(\mathbb{C})$ tensor products II: Puzzles determine facets of the Littlewood-Richardson cone," Journal of the American Mathematical Society **17**, 19–48 (2004).

[109] J. Watrous, "Semidefinite programs for completely bounded norms," Theory of Computing **5**, 217–238 (2009), arXiv:0901.4709.